

Improve accuracy of tracing production qualities using successive reconciliation

Use this technique for crude oil composition tracing and gasoline blend component quality prediction

J. D. KELLY and J. L. MANN, Honeywell Process Solutions, Toronto, Ontario, Canada; and F. G. SCHULZ, Willowend Technology, Houston, Texas

A relatively straightforward technique to improve the accuracy of predicting or tracing the internal qualities inside a production network when all of the stream qualities cannot be explicitly measured is presented. The qualities usually represent intensive variables such as compositions, properties and conditions.

Compositions can be the volume fractions of certain crude oil species in a crude oil mixture and usually the sum of the compositions equals unity or some other known constant. Properties can be the density, molecular weight or fat content of a stream. Conditions are the physical state attributes of a stream such as its temperature, pressure or velocity. Properties can also be functions of both its stream's compositions. Conditions such as density of a stream can alternatively be calculated by the sum of its compositions times the individual species density. The stream quantities, which also include the inventory or holdup amounts for equipment that can support accumulation, are the extensive variables of the system and most often have more measurements available due to their lower cost (i.e., an orifice plate with a differential pressure transmitter is considerably less expensive than a gas chromatograph with integrator).

Qualities of a production system are inferred over some tracing time horizon using discretized time by using known time periods or time slices. By assuming that the quantity values are error or defect free and then using well-known quantity and quality material balances at each time period, the dependent qualities starting from known initial or starting values can be calculated.

Unfortunately, the proliferation or propagation of quantity gross errors can seriously distort accuracy of these quality estimates even if accurate quality measurements exist. One remedy is to perform quantity and quality reconciliation at each time period in the tracing horizon, starting from the first time period. This enables the analyst performing the task of tracing over some horizon of, say, a day to detect and isolate defects in both time and space where space implies some location in the production network.

If an error is observed that is statistically significant, that measurement can be removed from the information set (i.e., deleted from the balance equations) and calculated provided enough other information exists. That is, if the measurement is redundant then the calculated value is observable.¹

The two sections to follow highlight the approach and its value. However, we briefly elucidate the business significance of tracing qualities in a production network. If all of the feedstocks, intermediates (work-in-progress) and products were known pure components or species, in essence only its material name and quantity would be required to manage the production in terms of meeting the plant's necessary supply and demand orders. This is typically the case in discrete-parts manufacturing, but it is more the exception than the rule in the process industries, especially in petroleum refining where a crude oil by its very nature is a collection of almost unknown proportions and properties.

Due to side-reactions, nonideal separations, degrading performance and transition cycles, the quality of the materials flowing through any process industry plant can also vary dramatically from one time to the next. Thus, careful attention must be paid to the tracing of qualities throughout the production timeline and network so that accurate and timely information is available when making business decisions. These may include deciding at what time to sell a particular material when its quality meets, but does not exceed, the customer's specification limits.

Successively reconcile quantities and qualities simultaneously.

Reconciling production data over a finite time horizon, minimizing the sum of squares of measured minus reconciled values and discretizing at uniformly or even non-uniformly spaced time-periods is not new and can be found in, for example, Albuquerque and Biegler² and Binder, et al.³ The value of this approach isn't solving successively over the multiple time periods comprising the horizon, but to rigorously include nonlinear data reconciliation to the business problem of tracing accurately production qualities. As mentioned previously, current practice is to assume the quantity values such as stream flows and inventories as error free and to calculate the dependent quality variables. Notwithstanding, this technique can improperly estimate the qualities when either measurement or modeling defects occur.

Model mismatch is one source of error. However, this is very difficult to isolate given that it is most often confounded by the measurement errors, unless explicit and usually intrusive perturbations (i.e., dithering signals) are injected as is used in advanced

INFORMATION TECHNOLOGY

TABLE 1. Quantity and quality names and measurement types.

Quantity name	Measurement type	Quality name	Measurement type
F1.flow	Measured	F1.C1, C2	Measured
F2.flow	Measured	F2.C1, C2	Unmeasured
F3.flow	Measured	F3.C1, C2	Unmeasured
F4.flow	Unmeasured	F4.C1, C2	Unmeasured
F5.flow	Measured	F5.C1, C2	Unmeasured
F6.flow	Measured	F6.C1, C2	Measured
T1.openInv	Fixed	T1.openinv.C1, C2	Fixed
T1.closeInv	Measured	T1.closeinv.C1, C2	Unmeasured
T2.openInv	Fixed	T2.openinv.C1, C2	Fixed
T2.closeInv	Measured	T2.closeinv.C1, C2	Unmeasured
T3.openInv	Fixed	T3.openinv.C1, C2	Fixed
T3.closeInv	Measured	T3.closeinv.C1, C2	Unmeasured

process control for closed-loop identification. The approach described here is to use steady-state data reconciliation solving techniques^{4,5,6,7} and to solve, one time period after the other, for the qualities—also reconciling any measured quantities and measured qualities if these are provided.

The steady-state data reconciliation solution required is nonlinear due to, at minimum, the product of quantity times quality variables found in the material balances. More complex nonlinearities are also possible. These usually involve constants and parameters that may need to be recalibrated using some form of designed or controlled experimentation. These more complicated relationships can also make the problem size and structure more troublesome during convergence, which may require the analyst to scale back the complexity.

When only the fundamental laws of conservation are employed (i.e., material, energy and momentum), auxiliary information is kept to a minimum and usually results in multilinear such as bilinear systems.⁸ The accumulating nature of the processes being reconciled and traced, which is the hallmark of unsteady or dynamic systems, is handled in this approach simply by including the process unit which is in an unsteady state to have an inventory variable; this inventory may or may not be measured.

The basic steps of the successive reconciliation to trace the qualities through the production network are as follows, where we assume the network model is available including the necessary equations to model splitters and perfectly mixed tanks, for example. The first step is to define the start of the tracing horizon whereby the necessary opening inventories and inventory qualities are provided for inventory carrying units. These may be either measured or simply calculated from the previous day's tracing.

The second step is to loop or cycle through each of the time periods one at a time and note which ones have detectable gross errors as indicated by the global or objective function test statistic.¹ At each of the time periods, the closing inventories and inventory qualities from the previous time period are used as the starting values for the next time period. In addition, initial guesses for the

TABLE 2. Perfect measured quantity stream flows and compositions over the 24-hr horizon using 1-hr time periods.

Time period	F1.flow, m ³ /h	F1.C1 and F1.C2, fraction	F2.flow, m ³ /h	F3.flow, m ³ /h	F5.flow, m ³ /h	F6.flow, m ³ /h	F6.C1 and F6.C2, fraction
1	10	0.1000 0.9000	10	10	5	5	0.4774 0.5226
2	10	0.1000 0.9000	10	10	5	5	0.4693 0.5307
3	10	0.1000 0.9000	10	10	5	5	0.4599 0.5401
4	9.5	0.1000 0.9000	10.5	9.5	5.5	4.5	0.4472 0.5528
5	9.5	0.1000 0.9000	10.5	9.5	5.5	4.5	0.4310 0.5690
6	9.5	0.1000 0.9000	10.5	9.5	5.5	4.5	0.4125 0.5875
7	9.5	0.1000 0.9000	10.5	9.5	5.5	4.5	0.3928 0.6072
8	9.5	0.9000 0.1000	10.5	9.5	5.5	4.5	0.3812 0.6188
9	9.5	0.9000 0.1000	10.5	9.5	5.5	4.5	0.3798 0.6202
10	9.5	0.9000 0.1000	10.5	9.5	5.5	4.5	0.3873 0.6127
11	9.5	0.9000 0.1000	10.5	9.5	5.5	4.5	0.4020 0.5980
12	10.5	0.9000 0.1000	6.5	10.5	4.5	5.5	0.4227 0.5773
13	10.5	0.9000 0.1000	6.5	10.5	4.5	5.5	0.4459 0.5541
14	10.5	0.9000 0.1000	6.5	10.5	4.5	5.5	0.4704 0.5296
15	10.5	0.9000 0.1000	6.5	10.5	4.5	5.5	0.4951 0.5049
16	10.5	0.1000 0.9000	6.5	10.5	4.5	5.5	0.5137 0.4863
17	10.5	0.1000 0.9000	6.5	10.5	4.5	5.5	0.5247 0.4753
18	10.5	0.1000 0.9000	6.5	10.5	4.5	5.5	0.5285 0.4715
19	10.5	0.1000 0.9000	6.5	10.5	4.5	5.5	0.5263 0.4737
20	10.5	0.1000 0.9000	6.5	10.5	4.5	5.5	0.5195 0.4805
21	10.5	0.1000 0.9000	6.5	10.5	4.5	5.5	0.5092 0.4908
22	10.5	0.1000 0.9000	6.5	10.5	4.5	5.5	0.4966 0.5034
23	10.5	0.1000 0.9000	6.5	10.5	4.5	5.5	0.4823 0.5177
24	9.5	0.1000 0.9000	7.5	9.5	5.5	4.5	0.4681 0.5319

TABLE 3. Traced (reconciled) stream compositions using Table 2 data.

Time period	F2.C1 and F2.C2, fraction		F3.C1 and F3.C2, fraction		F4.C1 and F4.C2, fraction	
1	0.7002	0.2998	0.3756	0.6244	0.4774	0.5226
2	0.5004	0.4996	0.4328	0.5672	0.4693	0.5307
3	0.3671	0.6329	0.4179	0.5821	0.4599	0.5401
4	0.2812	0.7188	0.3868	0.6132	0.4472	0.5528
5	0.2210	0.7790	0.3553	0.6447	0.4310	0.5690
6	0.1793	0.8207	0.3269	0.6731	0.4125	0.5875
7	0.1511	0.8489	0.3024	0.6976	0.3928	0.6072
8	0.4302	0.5698	0.3289	0.6711	0.3812	0.6188
9	0.6126	0.3874	0.3732	0.6268	0.3798	0.6202
10	0.7289	0.2711	0.4209	0.5791	0.3873	0.6127
11	0.8013	0.1987	0.4659	0.5341	0.4020	0.5980
12	0.8475	0.1525	0.5036	0.4964	0.4227	0.5773
13	0.8711	0.1289	0.5378	0.4622	0.4459	0.5541
14	0.8836	0.1164	0.5682	0.4318	0.4704	0.5296
15	0.8904	0.1096	0.5951	0.4049	0.4951	0.5049
16	0.5773	0.4227	0.5900	0.4100	0.5137	0.4863
17	0.3951	0.6049	0.5702	0.4298	0.5247	0.4753
18	0.2864	0.7136	0.5445	0.4555	0.5285	0.4715
19	0.2201	0.7799	0.5171	0.4829	0.5263	0.4737
20	0.1788	0.8212	0.4902	0.5098	0.5195	0.4805
21	0.1526	0.8474	0.4647	0.5353	0.5092	0.4908
22	0.1356	0.8644	0.4410	0.5590	0.4966	0.5034
23	0.1245	0.8755	0.4193	0.5807	0.4823	0.5177
24	0.1176	0.8824	0.3978	0.6022	0.4681	0.5319

unmeasured variables can be taken as the previous time period's values. This can help with convergence and reduce the computation time.

The third step is to revisit the first time period with detected gross errors and to analyze the temporally isolated subsystem to further isolate and locate the defect using the well-known measurement and constraint test statistics.¹ Any measurements that are found to have statistically significant gross errors are set to unmeasured and the time period's reconciliation is rerun. All succeeding time periods must also be rerun to correctly propagate the modification to the network. The process repeats until no time periods with detectable gross errors exist. To illustrate the procedure, we next describe a numerical example that highlights the workflow.

Note that an alternative to stepping through each time period starting from the first time period and proceeding in sequence is to solve the entire multiple time period problem in one simultaneous optimization run. This can be achieved using commercial nonlinear programming codes.⁶ However, the optimization problem is much larger and more difficult to solve. The other disadvantage is the intimacy with the model and data at each time period in terms of the effectiveness of the defect detection and diagnosis.

Using successive single time period reconciliation, it is possible to determine the exact time period when an inconsistency is statistically observed. In the simultaneous approach, previous or past time periods can be used to avoid detecting gross errors in future time periods, which is technically correct but may confound the analysis and make the results less transparent during the investigation. Therefore, we recommend the successive recon-

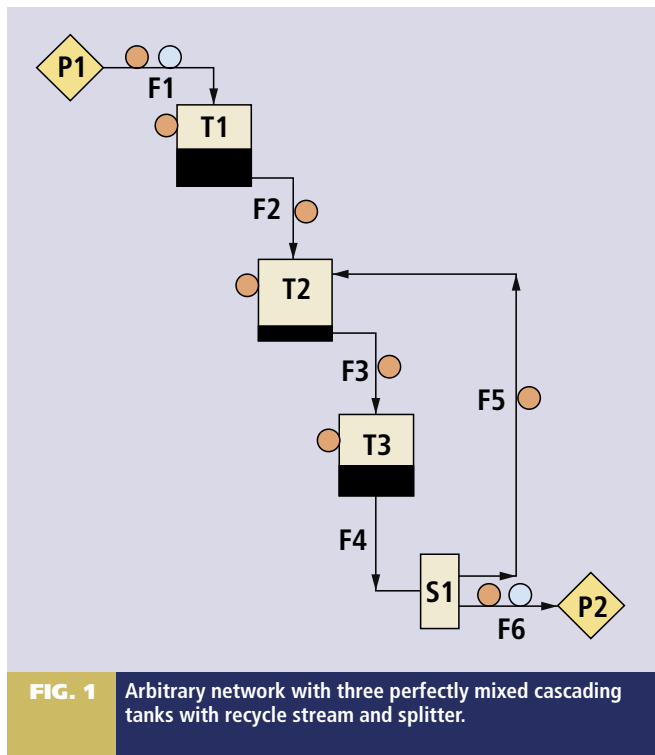


FIG. 1 Arbitrary network with three perfectly mixed cascading tanks with recycle stream and splitter.

ciliation approach to trace production qualities, although in the future combining both techniques may prove beneficial in some cases. Moreover, the successive reconciliation method can be used to provide valuable unmeasured variable starting values for the simultaneous multiperiod reconciliation, if used, which will aid in its convergence and robustness.

Example. A small but representative example details the effects of gross errors or defects in the measurement system. Fig. 1 shows the production network with hypothetical upstream and downstream process units, P1 and P2, as shown by the diamond objects. These process unit streams could also represent a pipeline or any other type of unit that surrounds the subsystem to be reconciled and traced.

Three perfectly mixed cascading tanks flow in series starting from tank T1. There are six stream flows, all of which are measured except for stream F4. The dark circles indicate flow measuring devices on the streams. A splitter, S1, splits stream F4 into streams F5 and F6. Stream F5 is a recycle stream that flows back to tank T2. Stream F6 is the flow that charges process unit P2. The two light circles on streams F1 and F6 represent analyzers that are able to measure the compositions of C1 and C2 (i.e., a gas chromatograph). These two compositions or concentrations also sum up to unity in all streams and are enforced by a normalization equality (i.e., $F6.C1 + F6.C2 = 1.0$, for example). Table 1 provides a summary of the quantity and quality system variables and available measurements.

In this example we provide quantity and quality measurements as found in Table 2 that constitute our perfect or error-free measurement values; no white noise is added. Our reconciliation and tracing horizon is 24 hr discretized into 1-hr time periods. Obviously, shorter time periods within the day are possible, but 1 hr is chosen for clarity of the presentation. The opening inventories for time period 0 (or the start of the horizon) are 10.0 m³, 20.0 m³ and 45.0 m³ for tanks T1, T2 and

INFORMATION TECHNOLOGY

TABLE 4. Deviation from Table 3's perfect stream compositions when gross errors are injected and reconciliation is performed.

Time period	F2.C1 and F2.C2, % diff.	F3.C1 and F3.C2, % diff.	F4.C1 and F4.C2, % diff.	Objective function value
1	-0.0346 0.0808	-0.0267 0.0161	-0.0038 0.0035	0.0001
2	-0.0739 0.0740	-0.0419 0.0320	-0.0102 0.0090	0.0007
3	-0.1174 0.0681	-0.0571 0.0410	-0.0179 0.0153	0.0022
4	0.9717 -0.3802	1.9959 -1.2591	0.3291 -0.2662	956.3
5	4.4205 -1.2539	3.6824 -2.0293	0.8667 -0.6566	2,717.4
6	8.4045 -1.8367	5.0532 -2.4542	1.4908 -1.0469	4,506.8
7	11.6068 -2.0653	6.1255 -2.6550	2.1341 -1.3805	5,989.4
8	-14.1088 10.6541	1.8082 -0.8861	2.0633 -1.2712	7,009.9
9	-15.3209 24.2229	-2.1764 1.2960	1.2873 -0.7882	7,558.2
10	-13.5525 36.4422	-4.5709 3.3219	0.1284 -0.0812	7,719.6
11	-11.0056 44.3862	-5.6833 4.9567	-1.0674 0.7175	7,601.2
12	-6.3589 35.3484	-7.2046 7.3081	-2.4756 1.8125	5,411.3
13	-4.5976 31.0734	-7.7921 9.0655	-3.6699 2.9535	4,283.9
14	-3.5949 27.2887	-7.8831 10.3733	-4.5723 4.0608	4,065.6
15	-2.8981 23.5533	-7.7034 11.3229	-5.1909 5.0899	4,230.6
16	18.1704 -24.8174	-5.6164 8.0828	-5.1743 5.4663	4,524.1
17	40.6438 -26.5497	-2.9389 3.8996	-4.6147 5.0946	4,851.0
18	62.2174 -24.9760	-0.1839 0.2198	-3.6763 4.1212	5,192.5
19	80.4170 -22.6986	2.4223 -2.5940	-2.5026 2.7809	5,560.0
20	93.4187 -20.3421	4.7796 -4.5957	-1.2019 1.2995	5,965.7
21	100.6128 -18.1160	6.8496 -5.9465	0.1472 -0.1527	6,414.0
22	102.5613 -16.0936	8.6267 -6.8067	1.4892 -1.4689	6,901.1
23	100.5035 -14.2924	10.1218 -7.3071	2.7858 -2.5955	7,419.1
24	96.5001 -12.8591	11.2863 -7.4548	3.8851 -3.4191	7,157.5

TABLE 5. Deviation when stream F2 and tank T1 quantities are declared as unmeasured and reconciliation is rerun.

Time period	F2.C1 and F2.C2, % diff.	F3.C1 and F3.C2, % diff.	F4.C1 and F4.C2, % diff.	Objective function value
1	-0.0346 0.0808	-0.0264 0.0159	-0.0038 0.0035	0.0001
2	-0.0739 0.0740	-0.0418 0.0319	-0.0102 0.0090	0.0007
3	-0.1176 0.0682	-0.0570 0.0409	-0.0179 0.0152	0.0022
4	-0.1592 0.0623	-0.0709 0.0447	-0.0259 0.0209	0.0044
5	-0.2043 0.0579	-0.0836 0.0461	-0.0342 0.0259	0.0074
6	-0.2510 0.0549	-0.0956 0.0464	-0.0428 0.0300	0.0110
7	-0.2970 0.0529	-0.1072 0.0465	-0.0516 0.0334	0.0152
8	-0.1114 0.0841	-0.1014 0.0497	-0.0594 0.0366	0.0195
9	-0.0737 0.1165	-0.0889 0.0529	-0.0647 0.0396	0.0234
10	-0.0559 0.1503	-0.0778 0.0566	-0.0672 0.0425	0.0260
11	-0.0458 0.1847	-0.0695 0.0607	-0.0676 0.0454	0.0274
12	-0.0404 0.2248	-0.0635 0.0644	-0.0664 0.0486	0.0281
13	-0.0358 0.2420	-0.0595 0.0692	-0.0645 0.0519	0.0284
14	-0.0315 0.2393	-0.0559 0.0735	-0.0623 0.0554	0.0287
15	-0.0275 0.2233	-0.0525 0.0772	-0.0600 0.0588	0.0291
16	-0.0650 0.0888	-0.0535 0.0770	-0.0586 0.0619	0.0292
17	-0.0763 0.0499	-0.0531 0.0705	-0.0575 0.0635	0.0297
18	-0.0773 0.0310	-0.0530 0.0633	-0.0567 0.0635	0.0296
19	-0.0763 0.0215	-0.0534 0.0571	-0.0560 0.0622	0.0288
20	-0.0770 0.0168	-0.0542 0.0521	-0.0556 0.0601	0.0276
21	-0.0800 0.0144	-0.0553 0.0480	-0.0554 0.0575	0.0262
22	-0.0849 0.0133	-0.0565 0.0446	-0.0555 0.0547	0.0249
23	-0.0905 0.0129	-0.0577 0.0417	-0.0557 0.0519	0.0237
24	-0.0936 0.0125	-0.0594 0.0392	-0.0562 0.0494	0.0228

T3, respectively. The corresponding start of horizon tank compositions for *C1* and *C2* are (1.0,0.0), (0.0,1.0) and (0.5,0.5), respectively, in each tank.

Using the perfect flow, inventory and composition values presented, the traced compositions for streams F2, F3 and F4 are shown in Table 3. These compositions are also the compositions for the material contained inside the tanks due to the perfectly mixed model, where compositions for streams F5 and F6 are identical to stream F4. When no errors exist in the data, reconciliation will produce identical results as simply tracing the qualities using fixed quantities. Required by the reconciliation to weight the sum of squares of adjustments objective function, we assume the standard deviation for both quantities and qualities is 1% of their measured value.

Now for demonstration purposes, we inject two gross errors into the quantities at different times. The first is located at time period 4 on stream F2 where we subtract 3.0 m³ from its true value, and this is sustained over the 24-hr horizon. The second is at time period 12 where we assume the inventory measurement on tank T1 follows the function $T1.closeinv(t) = (1 + 0.1t) T1.closeinv(t)$ where *t* is the time period index. Table 4 details the percent deviation from the perfect stream composition values when reconciliation is performed for each time period in succession. The objective function is also displayed, which has a threshold value of 12.6 assuming 95% confidence in the measurement set.

Clearly from Table 4, the gross errors in the quantities have severely distorted and corrupted the unmeasured stream compositions leaving the tanks by amounts greater than the precision of the analyzer (i.e., 1% of composition value). Fortunately, the reconciliation objective function starting from time period 4 is flagged as being statistically out-of-control when compared to the critical value of 12.6 that correctly detects the timing of the faults. The largest measurement statistic for time period 4 is 31.0 for the flow on stream F2, which is greater than its 95% confidence threshold value of 3.2—correctly indicating a strong gross error. At time period 12, the largest measurement error statistic is 60.3 for the closing inventory on tank T1, whereby we have also correctly isolated the second defect in the data set.

The next step is to declare the two quantities as unmeasured in the reconciliation. This effectively deletes the two quantity measurements from the system. Auspiciously, given the available redundancy in the network, the two deleted quantity variables are observable (i.e., the measurements are redundant). Table 5 shows the percent difference results including the objective function.

Visibly the percent deviation from the perfect stream compositions found in Table 5 are dramatically improved with no gross errors detectable given the near-zero objective function values. As the results of Table 4 indicate, when tracing alone is performed without reconciliation, potential gross errors can enter the system and, hence, skew the prediction of any unmeasured production qualities even when some measurements are available for the qualities. Yet when reconciliation and tracing are performed, valuable gross error statistics can be used to effectively remove infected measurements.

From the perspective of the computation time needed to solve one time period's reconciliation, it took on average 0.08 seconds and required typically four iterations on a Pentium III 500 MHz laptop using commercial software. There are 24 equations and 23 bilinear terms in the model with 8 quantity measured variables, 5 quality measured variables, 1 unmeasured quantity variable and 17 unmeasured quality variables.

This approach is also able to handle situations where the quality measurements arrive intermittently over the horizon when, for example, a laboratory sample result becomes available given that a dedicated field analyzer would be too expensive to install or unreliable and difficult to maintain. **HP**

LITERATURE CITED

- ¹ Crowe, C. M., "Data reconciliation—progress and challenges," *J. Proc. Cont.*, 6, pp. 89–98, 1996.
- ² Albuquerque, J. S. and L. T. Biegler, "Data Reconciliation and Gross-Error Detection for Dynamic Systems," *AIChE Journal*, 42, pp. 2,841–2,856, 1996.
- ³ Binder, T., L. Blank, W. Dahmen and W. Marquardt, "Regularization of dynamic data reconciliation problems by projection," IFAC Symposium, ADCHEM 2000, Pisa, Italy, June 2000.
- ⁴ Leibovici, C. F., V. S. Verneuil and P. Yang, "Improve prediction with data reconciliation," *Hydrocarbon Processing*, October 1993.
- ⁵ Kelly, J. D., "A regularized solution to the reconciliation of constrained data sets," *Computers Chem. Engng.*, 22, pp. 1,771–1,788, 1998.
- ⁶ Kelly, J. D., "Reconciliation of process data using other projection matrices," *Computers Chem. Engng.*, 23, pp. 785–789, 1999.
- ⁷ Poku, M. Y. B., L. T. Biegler and J. D. Kelly, "Nonlinear process optimization with many degrees of freedom in process engineering," *Ind. Eng. Chem. Res.*, 42, pp. 6,803–6,812, 2004.
- ⁸ Kelly, J. D., "Formulating large-scale quantity-quality bilinear data reconciliation problems," *Computers Chem. Engng.*, 28, pp. 357–366, 2004.



J. D. Kelly is a chemical engineer and has a master's degree in advanced process control from McMaster University. He has worked as an advanced control engineer at both Shell Canada and Imperial Oil, including implementing real-time optimization programs and tactical planning and scheduling solutions in their refineries. Mr. Kelly has installed plant-wide data reconciliation packages in several oil refineries around the world and has written many academic publications on the subject. He is now a solutions architect for advanced planning and scheduling at Honeywell Process Solutions in Toronto, Ontario, Canada.



J. L. Mann is a chemical engineer with a bachelor of applied science degree from the University of Toronto. He has worked as a design engineer and as a simulation engineer at Imperial Oil, including developing refinery-wide simulation tools to support planning and scheduling activities. Mr. Mann has worked on a number of plant information system projects with a focus on integrating plant data collection systems to plant-wide yield accounting systems. He now is a business architect for Honeywell Process Solutions in Toronto, Ontario, Canada.



Fred Schultz is a chemical engineer and software developer with an MS degree from the University of Houston. Currently principal consultant for Willowend Technology, Inc., he has 20 years of experience developing and deploying advanced solutions in the process industries.