Reconcile Quantity and Time to Diagnose Production Defects in the Process Industries

Jeffrey D. Kelly[1], John L. Mann[2] and Douglas G. Moffat[3]

June 2003

Keywords: Reconciliation, scheduling, quantity sizing, timing, logistics, multi-period, multi-purpose and meta-physical.

---

[1] Honeywell Industry Solutions, 300 Yorkland Blvd., Toronto, Ontario, Canada, M2J 1S1
E-mail: jeff.kelly@honeywell.com
[2] E-mail: john.mann@honeywell.com
[3] E-mail: doug.moffat@honeywell.com

Abstract

The purpose of this article is to express a novel stewardship procedure to aid in the detection and identification of production execution and process instrumentation problems. The technique bridges the gap between well-known data reconciliation and production scheduling by formulating both quantity and timing constraints into two reconciliation problems to be described. The method requires the use of two quadratic programs to statistically adjust flows and inventories and then to adjust start and end-times of actual production activities in order to help expose and isolate *mis-specified, mis-logged* and *mis-operated* production anomalies in the data. The value of such a procedure is to allow the production analyst to better understand and steward the performance, competency and efficiency of the manufacturing facility with respect to how well the plant is operated on a daily-basis. A poorly operated plant typically has higher direct manufacturing costs for its products, has lower customer service levels (less-reliable and less-predictable), has attenuated throughputs, is more prone to safety-related issues and is less environmentally sound leading to a higher overall operating liability. We demonstrate the technique using a small but representative illustrative example.

Introduction

Manufacturing industries are primarily divided into discrete-parts manufacturing and process industries. Process manufacturing is defined as production that adds value to materials by mixing, splitting, separating, transforming or chemical reaction (Crama *et. al.* (2001)). Process industries can be furthered decomposed into batch and continuous processes and are easily distinguishable. Batch processing is defined as a manufacturing technique in which components are *accumulated* and processed together in a group, consignment, lot or batch. Continuous processes do not accumulate the components before or during processing. The components are processed immediately upon contact with one another and flow continuously through the equipment according to a specific time rate of change of quantity and without significant hold-up. Though a finite amount of residence-time can be achieved in continuous processes by varying the flowrate, length and effective diameter of the process-side equipment, this residence-time or space-velocity is not the same as a batch, processing or cycle-time used to describe batch processes. A tank is a simple example of a batch process unit and a header is a trivial example of a continuous process unit. Semi-batch processes have some continuous process characteristics such as a continuous flow of a component over the batch-time and semi-continuous processes have some attributes of batch processes such as they can operate intermittently.

That said, our quantity and time reconciliation procedure is designed to detect and identify possible production anomalies and can be applied easily to both batch and continuous processes including any combinations thereof. Most process industry plants are what are known as *closed-shops* meaning that they require the activities, tasks, jobs, operations, actions or steps in the production to be *sized* in terms of quantity (Graves (1981)); the exception would be the packaging and bundling portions of their plants. *Open-shops* are typical of discrete-parts manufacturing processes and require a different type of technology in terms of how they are scheduled and hence how they are operated (Pinedo (1995)). Batch processes are sized by a batch-size expressed as a volume or weight amount (Maravelias and Grossmann, (2002)) and continuous processes are sized by a charge-size which is usually expressed as a volume or weight-based throughput or flowrate (Kelly (2002)). All process industry plants contain some form of finite inventory depending on the physical state of the material (i.e., gas, liquid, solid, slurry, etc.) where inventory vessels are shared and limited resources and need to be properly managed. It is the presence of material, inventory or quantity balances that fundamentally separates the process industries from the discrete-parts industries and classifies it as a closed-shop.

We also can make the statement that production in the process industries is comprised of three orthogonal dimensions and they are *quantity*, *logic* and *quality*. The quantities are the extensive flows and inventories, logic includes detailed timing and operating rules and polices, and quality is the degree of separation from conformance to product specifications such as intensive strictures on compositions, properties and conditions. It is also known that the combined quantity and logic problem is a *logistics problem* and the combined quantity and quality problem is a *quality problem* (Kelly and Mann (2003)). We mention this to draw attention to the fact that this study is fundamentally a logistics reconciliation method and is a relatively new idea. The non-linear quality reconciliation problem has been treated in the data reconciliation literature as can be found in Crowe (1996).

This work also touches on the philosophy of the quality management revolution such as those found in Six Sigma, TQM, World-Class Manufacturing, Kaizen, DMAIC, Deming Wheel, Shewhart Cycle and Continuous Improvement. All of these programs at the core exemplify the notion of a perpetual cycle of *plan-perform-perfect*. The main theme of this paper is very similar although we refer to it as *schedule-execute-examine* (SEE). The idea is to *schedule* the production based on the supply and demand-orders by generating production-orders, carry out or *execute* these production-orders and then to *examine* or analyze how well these orders were fulfilled. The SEE framework is intended to allow the production facility with the ability to “see” or have visibility into the future, present and past operations of the plant. This view into the full spectra of the temporal dimension of the production enables the production analyst to possibly recommend ways to *eliminate* defects after they have been detected and identified. Eliminating defects in a production or manufacturing system can occur through several avenues such as better communication, coordination and education of the operating staff, new technologies, preventative maintenance and perhaps a change in the philosophy at the plant by rewarding for improving performance as opposed to punishing for not meeting the production objectives.

Now that we have defined the industry scope of the opportunity, it is important to frame the discussion around a specific instance of the problem in order to adequately articulate the business value and need. We do this through a small illustrative example found in Figure 1.

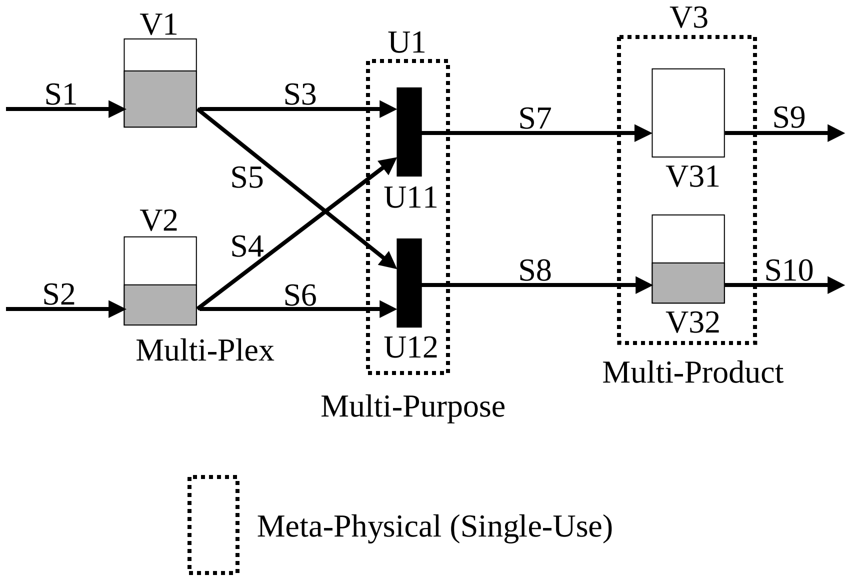

Figure 1. Closed-shop illustrative example flow-chart with two meta-physical objects.

This is closed-shop which involves three physical vessels or tanks V1, V2 and V3 where V3 is a swing or *multi-product* tank and can store two different types of materials but not at the same time as indicated by the rectangular dotted line box; we call this a single-use logic constraint. There is a physical unit U1 which is a semi-continuous unit and can operate intermittently when required to mix materials from tanks V1 and V2 in different and pre-specified proportions. It can produce two distinctly different products and is hypothetically modeled as two *meta-physical* units U11 and U12 as seen by the dotted rectangle surrounding it. There are ten streams connecting the physical and meta-physical equipment with the out flows of tanks V1 and V2 being *multi-plexed* between U11 and U12 where streams S3 and S5 (and S4 and S6) cannot both be active at the same time. Streams have individual moves attached to them when there are active or flowing. These moves are no different than an instance of an activity or task taking place and being supported by the stream. The operating rules or logic of the shop are found in Table 1.

Table 1. Logic rules for units, vessels and streams (N/A – not applicable).

Logic Rule	**V1**	**V2**	**U1**	**V3**	**Streams**
Semi-Continuous	N/A	N/A	Yes	N/A	All
Single-Use	N/A	N/A	Yes	Yes	N/A
Fill-Draw-Delay	0.0 h	No	N/A	1.0 h	N/A
Lower Up-Time	N/A	N/A	2.0 h	N/A	S7, S8

These rules along with capacity bounds on the streams and equipment would dictate how a production schedule would be constructed to meet the supply and demand requirements or orders of the process (i.e., S1, S2, S9 and S10 respectively). The *semi-continuous* attribute indicates which pieces of equipment are continuous in nature and not batch. Streams are almost always semi-continuous when dealing with multi-purpose equipment as when a source or destination unit/vessel is not active the stream is also not active by implication. The *single-use* or unary resource constraint, as mentioned, ensures that for those multi-functional equipment only one activity can occur at any given time. The *fill-draw-delay* rule means that after a fill activity or stream inlet move a draw or stream outlet move can only occur after a certain time-interval has elapsed. This is to support a sampling, mixing or certification delay on the vessel. A fill-draw-delay of 0.0 hours indicates that the vessel is designated as *standing-gauge* meaning that there cannot be flow in and out at the same time though there is no significant wait time before an opposite move can take place. The *lower up-time* restriction is identical to a minimum run-length constraint. This constraint enforces that any stream move activity if active must exist for a time greater than or equal to the lower up-time; there can also be an upper up-time. For this example streams S7 and S8 also imply the up-time of U1 so an explicit lower up-time on U1 is not required. However, in more generic unit operations a particular piece of equipment can also have an up-time especially if there are one or more outlet streams available. There can also be up-time constraints on the material service or material designation of the vessel in that if a switch-over to another material service occurs it must last for a specified number of hours; this is not considered in this example. There are also several other possible operating rules not mentioned but these can be found in Kelly (2002) and Kelly and Mann (2003).

From the perspective of differences between batch and continuous processes it is also interesting to note that batch process operation and scheduling is more *unit/vessel-based* in terms of how the production activities are assigned to equipment. In continuous processes the operation and scheduling is more *stream-based* as can be seen by the time-chart in Figure 2 for a sample thirty-two hour set of stream moves. If this were a batch process time-chart we would see for example V1, V2, U1 and V3 on the left-hand-side. The two outlooks are in fact identical from an information content point of view and are complementary to one another in terms of presentation and perception. Nevertheless, this should not distract from our discussion that the technique provided in this article can be applied equally well to batch, continuous and hybrid processes. The dot-line-dot objects indicate the move instances for each stream. The transparent dots at the bottom of the chart show the times when a production event occurs over all streams such as a start-up of any move on a stream or a shut-down of a move; these will form the boundaries or event points of the time-periods.

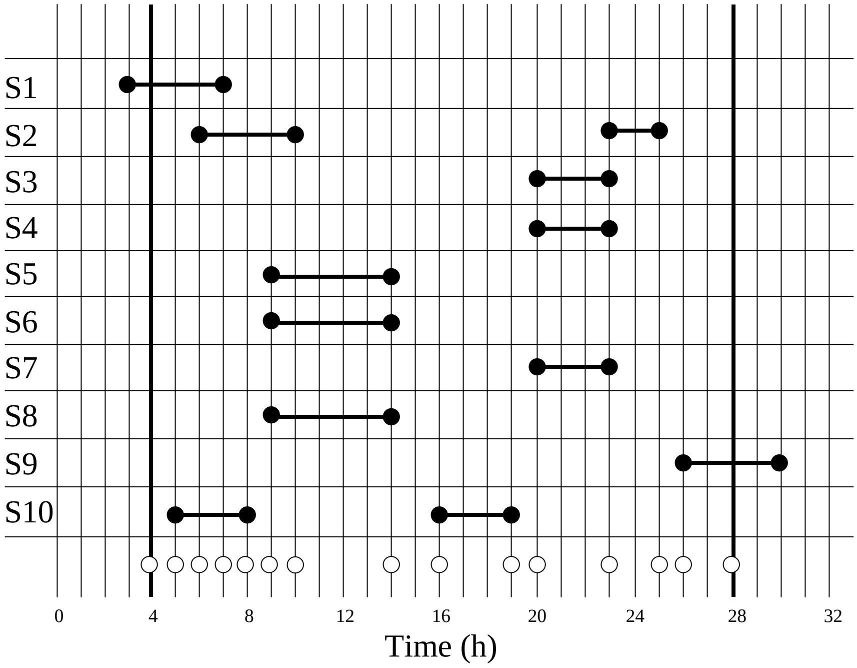

Figure 2. Illustrative example time-chart using streams as left-hand-side resources.

An important aspect of our approach is that we do not allow *re-sequencing* or *re-assigning* of the moves during the reconciliation. Given that all closed-shop production scheduling problems involve in some form or another sizing, sequencing, assigning and timetabling of schedulable activities, we restrict ourselves to only *re-sizing* and *re-timing* of the stream moves. This is a requirement because re-sequencing and re-assigning activities to resources necessitates a combinatorial search procedure such as branch-and-bound (Gueret *et. al.* (2002)). Such decision-making is known as disjunctive programming (Maravelias and Grossmann (2002)) given the either-or decisions of whether to start or stop a move to keep a vessel from over or under-flowing. Hence, if a stream move is logged or recorded as having actually happened, we assume that only its flow quantity and/or start and end timing is suspect. The stream move will exist through out the two reconciliation problems to be presented forthcoming. Another important assumption we make is that for any stream move the flowrate during the move is constant but usually unknown. There are flowrate lower and upper bounds included with each move where its actual stream move flowrate is not explicitly known or required to be known. This means that during a move activity, the flowrate is constant and equal to the quantity divided by the end-time minus the start-time (duration) of the move, subject to a lower and upper bound. Each stream has a lower and upper flowrate limit which usually does not change during the *reconciliation horizon* in question. This reconciliation horizon is usually about 24 hours and is broken down as in Figure 2 into time-periods according to the measured start and end-time events logged. Any horizon longer than a day will make it more difficult to troubleshoot the underlying root cause or fault of the observed effect or symptom.

At this moment it is appropriate to present the workflow of the procedure. Figure 3 diagrammatically describes the technique where we start by retrieving the measured, actual or logged stream movements from some transactional information system in the plant typically logged manually by the plant operators. Individual move records contain the stream identification (sometimes with source and destination equipment), start and end-time, the material being transferred and the quantity size of the move. Time-periods and time-period durations are then determined by shaking out the event times (so to speak) over all streams and placing them on a single timeline as in Figure 2. There is no requirement that the time-period durations be equal. Vessel inventory gauge and time-period sliced flowmeter readings are then retrieved from some real-time data historian at each time-period where for inventories, the value is taken to be at the end of the time-period. It is also possible to slice or pro-rate the stream move quantity sizes directly from the move records by multiplying the move size by the time-period duration divided by the total stream move duration. Either approach is acceptable depending on the level of precision and accuracy of the data.

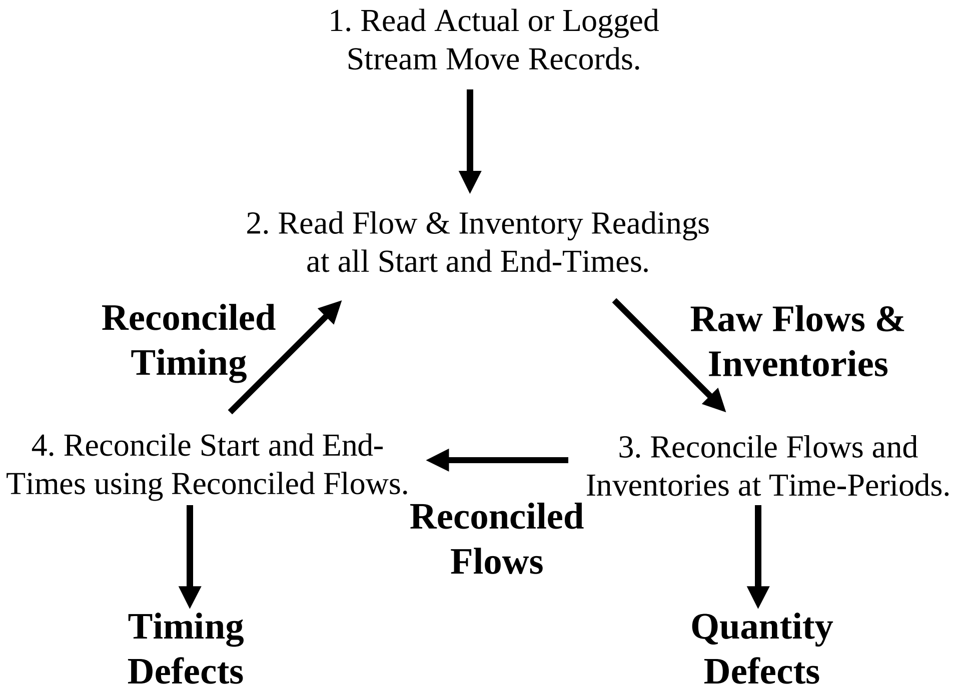

Figure 3. Overall workflow of the reconciliation procedure.

The next step is to build a quantity balance model for the reconciliation horizon in question and to set-up a quadratic program (QP) to solve for the adjusted flow and inventory values at the time-periods subject to flowrate and inventory lower and upper bounds. This quantity reconciliation will be described in more detail in the next section. Provided no gross errors are detected as indicated by the magnitude of the weighted sum of squares objective function, the reconciled time-period flows are re-constituted back into the stream move quantities and used in the time reconciliation QP described later. If this reconciliation is also free of significant gross errors then procedure can terminate. Else there maybe a requirement to use the reconciled times to re-slice the reconciliation horizon and to re-start the procedure in the hopes that more accurate quantities and timing will result. Notwithstanding, the major benefit of the technique is to expose and to uncover production irregularities such as mis-specified, mis-logged or mis-operated faults or defects in the production setting. If that can be achieved in only one pass, without using the reconciled timing, then it is acceptable to stop at that point.

Quantity Reconciliation

Steady-state data reconciliation is a well-studied discipline in the chemical engineering literature as surveyed by Crowe (1996) with data reconciliation applied to dynamic systems growing in importance and can be found in Albequerque and Biegler (1996), Bagajewicz and Jiang (1997) and Binder *et. al.* (2000). The quantity reconciliation performed in this study is a multi-period data reconciliation formulation which is in essence steady-state reconciliation applied over multiple time-periods to approximate a dynamic production system but solved simultaneously. Plant-wide reconciliation of quasi steady-state production data, typically performed using a one-day mono-period reconciliation horizon, is now common place in most oil-refineries and petrochemical plants (Kelly (1999) and (2000)). In these applications all of the stream movements are effectively stretched from the start to the end of the reconciliation horizon whereby all quantities and timing within the horizon being lost. One of the main focuses of this article is to try to extract more diagnostic information for the “within the horizon” events given that gross error or defect cancellation can occur quite readily due to the temporal accrual of the stream moves in and out of the processing units and storage vessels. By performing a multi-period quantity reconciliation we can limit or minimize the degree to which *Type* I and II errors occur, that is, false positives (phantom errors) and false negatives (missing errors) respectively. The basic assumption of data reconciliation assumes that a measurement is composed of three parts as shown in equation (1).

$$x = \mu + \varepsilon + \delta \quad (1)$$

where x is the measurement, μ is the mean, ε is a normally distributed random error with a measurement variance of σ^2 and δ is a gross error of arbitrary distribution. In the absence of δ the challenge of data reconciliation is to use the spatial redundancy in the system (i.e., the quantity balances) to more precisely estimate μ from the measurements. This approximation is called the reconciled value and is equal to the measurement minus the adjustment determined by the reconciliation solver. The injection or insertion of a defect into the measurement is most often completely unknown and can only be observed through larger than normal adjustments in the measurements. This leads us to our current opportunity and that is how to formulate a quantity reconciliation system to assist in the diagnosis of a good or bad production execution over typically a one-day reconciliation horizon.

In order to present the formulation in a meaningful manner we model the quantity reconciliation using the specific scenario depicted in Figure 1 and 2. The first relationship to define is the quantity reconciliation weighted sum of squared adjustments objective function J_Q.

$$J_Q = \sum_{t=1}^{NT=14} \left(\sum_{s=1}^{NS=10} w_s \cdot (f_{s,t}^{a})^2 + \sum_{v=1}^{NV=4} w_v \cdot (h_{v,t}^{a})^2 \right) + w_Q \cdot \left(\sum_{s=1}^{NS} \sum_{m=1}^{NM_s} (afr_{s,m}^{L})^2 + (afr_{s,m}^{U})^2 + \sum_{v=1}^{NV} \sum_{t=1}^{NT} (ah_{v,t}^{L})^2 + (ah_{v,t}^{U})^2 \right) \quad (2)$$

where $f_{s,t}^{a}$ is the sliced time-period flow adjustment for each stream at each time-period with w_s being the weighting parameter which could be temporally specified and is related to the stream flow measurement uncertainty or error. The inventory or hold-up adjustment $h_{v,t}^{a}$ for each vessel at each time-period is weighted by w_v where the second term in the equation multiplied by w_Q is described in more detail in equations (4) and (5). The number of streams NS and the number of vessels NV are also stated explicitly in the sums for the illustrative example. The number of time-periods NT and their location in time can be found in Figure 2 and is shown in Table 2 where the duration of each time-period is fixed for the measured start and end-times retrieved from the stream move records. The reconciliation horizon is determined by summing up all of the time-periods (see last row in table).

Table 2. Asynchronous time-periods for quantity reconciliation.

Time-Period	Start-Time (h)	End-Time (h)	Duration (h)
0	4.0	4.0	0.0
1	4.0	5.0	1.0
2	5.0	6.0	1.0
3	6.0	7.0	1.0
4	7.0	8.0	1.0
5	8.0	9.0	1.0
6	9.0	10.0	1.0
7	10.0	14.0	4.0
8	14.0	16.0	2.0
9	16.0	19.0	3.0
10	19.0	20.0	1.0
11	20.0	23.0	3.0
12	23.0	25.0	2.0
13	25.0	26.0	1.0
14	26.0	28.0	2.0
		Total	24.0

The individual stream move records (without material specification) can be found in Table 3 without any superimposed random errors (i.e., $\varepsilon = 0$) hence the use of the word nominal in the table caption. Table 4 shows them as stream flows after being sliced into the appropriate time-periods. Similarly the vessel nominal inventory profiles are displayed in Table 5 for each time-period where it is defined that the inventory value is actually at the end of the time-period (i.e., closing inventories for each period).

Table 3. Quantity sizes (nominal) of stream moves from start-time to end-time.

Stream, Move	Size (m3)	Rate (m3/h)	Start-Time (h)	End-Time (h)	Duration (h)
S1, 1	20.0	5.0	3.0	7.0	4.0
S2, 1	20.0	5.0	6.0	10.0	4.0
S2, 2	10.0	5.0	23.0	25.0	2.0
S3, 1	9.0	3.0	20.0	23.0	3.0

S4, 1	9.0	3.0	20.0	23.0	3.0
S5, 1	5.0	1.0	9.0	14.0	5.0
S6, 1	10.0	2.0	9.0	14.0	5.0
S7, 1	18.0	6.0	20.0	23.0	3.0
S8, 1	15.0	3.0	9.0	14.0	5.0
S9, 1	28.0	7.0	26.0	30.0	4.0
S10, 1	21.0	7.0	5.0	8.0	3.0
S10, 2	21.0	7.0	16.0	19.0	3.0

Table 4. Quantity sizes (nominal) for stream flows in time-periods.

Stream	Time-Period	Size (m3)	Stream, Move
S1	1	5.0	S1, 1
S1	2	5.0	S1, 1
S1	3	5.0	S1, 1
S2	3	5.0	S2, 1
S2	4	5.0	S2, 1
S2	5	5.0	S2, 1
S2	6	5.0	S2, 1
S2	12	10.0	S2, 2
S3	11	9.0	S3, 1
S4,	11	9.0	S4, 1
S5	6	1.0	S5, 1
S5	7	4.0	S5, 1
S6	6	2.0	S6, 1
S6	7	8.0	S6, 1
S7	11	18.0	S7, 1
S8	6	3.0	S8, 1
S8	7	12.0	S8, 1
S9	14	14.0	S9, 1
S10	2	7.0	S10, 1
S10	3	7.0	S10, 1
S10	4	7.0	S10, 1
S10	9	21.0	S10, 2

Table 5. Quantity sizes (nominal) for vessel inventories at end of each time-period.

Time-Period	V1 (m3)	V2 (m3)	V31 (m3)	V32 (m3)
0	10.0	5.0	0.0	27.0
1	15.0	5.0	0.0	27.0
2	20.0	5.0	0.0	20.0
3	25.0	10.0	0.0	13.0
4	25.0	15.0	0.0	6.0
5	25.0	20.0	0.0	6.0
6	24.0	23.0	0.0	9.0
7	20.0	15.0	0.0	21.0
8	20.0	15.0	0.0	21.0
9	20.0	15.0	0.0	0.0
10	20.0	15.0	0.0	0.0
11	11.0	6.0	18.0	0.0
12	11.0	16.0	18.0	0.0
13	11.0	16.0	18.0	0.0
14	11.0	16.0	4.0	0.0

The balance constraints that the problem is subject to are the material balances around each of the four vessels and the two units as follows

$$(h_{v,t-1}^m + h_{v,t-1}^a) + \sum_{s=1, s\in U2V}^{NS} (f_{s,t}^m + f_{s,t}^a) = (h_{v,t}^m + h_{v,t}^a) + \sum_{s=1, s\in V2U}^{NS} (f_{s,t}^m + f_{s,t}^a) \ \forall \ v = 1..NV, t = 1..NT \quad (3a)$$

$$\sum_{s=1, s\in V2U}^{NS} (f_{s,t}^m + f_{s,t}^a) = \sum_{s=1, s\in U2V}^{NS} (f_{s,t}^m + f_{s,t}^a) \ \forall \ v = 1..NU, t = 1..NT \quad (3b)$$

where inventory and flow variables with the superscript "m" are the measured values and $U2V$ and $V2U$ are the mapping matrices indicating which streams are connected from units to vessels and from vessels to units respectively. In this example, streams S1, S2, S7 and S8 can be considered as units to vessels and streams S3, S4, S5, S6, S9 and S10 are vessel to unit streams. We also impose flowrate constraints on each of the streams in each time-period as

$$\sum_{t=ST_{s,m}}^{ET_{s,m}} fr_s^L \cdot dt_t^m - \sum_{t=ST_{s,m}}^{ET_{s,m}} (f_{s,t}^m + f_{s,t}^a) \le afr_{s,m}^L \ \forall \ s = 1..NS, m = 1..NM_s \quad (4a)$$

$$\sum_{t=ST_{s,m}}^{ET_{s,m}} (f_{s,t}^m + f_{s,t}^a) - \sum_{t=ST_{s,m}}^{ET_{s,m}} fr_s^U \cdot dt_t^m \le afr_{s,m}^U \quad (4b)$$

where NM_s specifies the number of moves for each stream over the reconciliation horizon, dt_t^m is the duration of each time-period found in column four of Table 2, fr_s^L and fr_s^U are the flowrate lower and upper bounds shown in Table 6 and $ST_{s,m}$ and $ET_{s,m}$ are the starting and ending time-period indices for each stream move found in Table 4. The artificial, penalty, elastic or alarm variables $afr_{s,m}^L$ and $afr_{s,m}^U$ are minimized in J_Q along with the adjustments and are weighted strongly by the parameter w_Q found in equation (2). If any of these are non-zero at the solution then the flow and inventory adjustments could be changed to respect equations (4a) and (4b). Consequently, in order not to make the QP go hard infeasible the flowrate bounds are dualized and are made soft in the objective function. Thus, if any of these artificial variables are non-zero at the solution then this explicitly implies that gross errors are present in the system and further scrutiny is required to ascertain a plausible cause.

Table 6. Lower and upper flowrate bounds for streams.

Stream	Lower Rate (m3/h)	Upper Rate (m3/h)
S1	4.0	6.0
S2	4.0	6.0
S3	2.0	4.0
S4	2.0	4.0
S5	1.0	2.0
S6	1.0	3.0
S7	5.0	7.0
S8	2.0	4.0

S9	6.0	8.0
S10	6.0	8.0

The final set of bounds in the formulation, with artificial variables, are for the vessel lower and upper inventory capacities arbitrarily set for our example at 0 m3 and 50 m3 respectively.

$$h_v^L - (h_{v,t}^m + h_{v,t}^a) \le ah_{v,t}^L \;\; \forall \;\; v = 1..NV, t = 1..NT \qquad \text{(5a)}$$

$$(h_{v,t}^m + h_{v,t}^a) - h_v^U \le ah_{v,t}^U \qquad \text{(5b)}$$

where $ah_{v,t}^L$ and $ah_{v,t}^U$ are the artificial variables for each vessel, for each time-period weighted strongly in equation (2). These constraints can be neglected whereby any under or over flow can be calculated after the reconciliation solution. If a vessel exceeds its capacity bounds then a symptom of a fault has been detected and should be investigated.

After the solution of the quantity reconciliation, the results must analyzed cautiously for gross errors before the time reconciliation step can be performed. If the quantity reconciliation is deemed to be absent of significant faults or defects then the reconciled flow adjustments $f_{s,t}^a$ and the time sliced measured flows $f_{s,t}^m$ can be summed together to provide a better estimate of the flow quantities for the stream moves found in Table 3.

It should also be pointed out that formulating a multi-period quantity reconciliation is not new and can be found in Fillon *et. al.* (1995). These authors create a multi-period problem for an experimental batch process and solve it using conventional steady-state linear data reconciliation techniques without inequalities (i.e., without the flowrate and inventory bounds). In order to solve their batch process reconciliation problems they were required to extend the reconciliation formulation to multi-period as is done in this work to support either batch, continuous or hybrid systems. In terms of more detailed gross error detection and identification techniques, the reader should also be referred to the work of Crowe (1988) and Harikumar and Narasimhan (1993). Crowe showed that the absolute value of the maximum-power gross error statistics are equivalent to the square root of the change in the quadratic objective function when the measurement in question is deleted or made *unmeasured* in the reconciliation problem. This is useful because to determine the maximum-power statistics requires the covariance matrix of the adjustments to flow and inventory to be known which is complicated by the presence of inequalities and penalties (i.e., equations (4) and (5)). However, Harikumar and Narasimhan show that active inequalities at the QP solution can be made equalities whereby a conventional linear data reconciliation solver can be executed to derive the covariance matrix of the adjustments with these extra balances added. Unfortunately, this requires another solver beyond the QP to compute the covariance matrix. A simpler solution, albeit more computationally intensive, based on the observation of Crowe is to run the QP as many times as there are measurements to compute the change in the objective function when one-at-a-time deletions are made for each measurement. This will accurately compute the maximum-power statistics which are the most powerful indicators of gross errors when only one gross error exists (Crowe (1988)). It can be furthered improved from an efficiency point of view by only deleting measurements with significant adjustments.

Time Reconciliation

Reconciling or adjusting the start and end-times for each stream move instance is the focus of the time reconciliation QP. The underlying time balances or constraints are more complicated than those presented in the quantity reconciliation because we now need to deal with operating procedures or logic transcribed as timing restrictions. These are presented in the order found in Table 1 but first we describe J_T the quadratic objective function.

$$J_T = \sum_{s=1}^{NS} \sum_{m=1}^{NM_s} \left(w_{st,s} \cdot (st^a_{s,m})^2 + w_{et,s} \cdot (et^a_{s,m})^2 \right) + \ w_T \cdot TimeArtificials^2 \quad (6)$$

where $st^a_{s,m}$ and $et^a_{s,m}$ are the adjustments to the stream move start and end-times with corresponding weights shown. The second term containing $TimeArtificials^2$ is the cumulative sum of all the artificial variables to be individually described in the constraints to follow where w_T of course is the overall penalty weighting for the timing artificials. The first set of timing constraints deal with the semi-continuous nature of each stream move as

$$st^m_{s,m} + st^a_{s,m} - (et^m_{s,m} + et^a_{s,m}) \le asc_{s,m} \ \ \forall \ \ s = 1..NS, m = 1..NM_s \qquad (7)$$

where $asc_{s,m}$ is a time artificial variable and is summed and squared in equation (6). This inequality states that all streams must have a move duration greater than or equal to zero else a non-zero artificial variable will result implying the move is mis-specified. Semi-continuous constraints on the units is more complicated in that there needs to be a time-ordering or time-sequencing of the stream move records before generation of these constraints in order to properly represent the stream move sequence.

$$\begin{aligned} st^m_{s,m} + st^a_{s,m} - (st^m_{s',m'} + st^a_{s',m'}) = asc^+_{u,mu} - asc^-_{u,mu} \ \forall \ \mathrm{u} = 1..NU, s, s' = 1..NS, mu = 1..NM_u, \\ m = 1..NM_s, s \in V2U, m' = 1..NM_{s'}, s' \in U2V \end{aligned} \qquad (8a)$$

$$et^m_{s,m} + et^a_{s,m} - (et^m_{s',m'} + et^a_{s',m'}) = asc^+_{u,mu} - asc^-_{u,mu} \qquad (8b)$$

where s' and m' are the representative or reference stream and move indices for the stream moves on the outlet side of the unit and $asc^+_{u,mu}$ and $asc^-_{u,mu}$ are the artificial variables for semi-continuous unit indexed by the unit number and the move number on the unit. There are two of them given that all of the artificial variables are declared to be non-negative although this is more of a preference than a true requirement. The number of moves on a semi-continuous unit is indicated by NM_u which is easily calculated as the number of unique move sets on the unit given the stream move in's and stream move out's over the reconciliation horizon. For the illustrative example $NM_{U11} = 1$ and $NM_{U12} = 1$. It is also important to note that if there was more than one stream out of each unit then it would be more efficient in terms of the number of inequalities generated to choose only one of the stream move out's as the reference stream move (i.e., s' and m'). Single-use constraints are really only appropriate for physical units and vessels (i.e., U1 and V3) and are modeled in equations (9a) and (9b). The actual stream move records must be properly time-ordered in order to properly reference the correct stream move pair. For our illustrative example there are only one stream out's for each single-use units and vessels (i.e., S7 and S8 and S9 and S10 respectively) which makes the setup of the constraints easier than if there were more than one stream out. When there is only one stream out it is easy to use the one outlet stream as the indicator for single-use; the same would be true for one stream in.

$$\begin{aligned} et^m_{s,m} + et^a_{s,m} - (st^m_{s',m'} + st^a_{s',m'}) \le asu_{u'',mu} \quad \forall\ u'' \in SUU, mu = 1..NM_{u''}, s, s' \in S_{u''}, \\ m = 1..NM_s, m' = 1..NM_{s'}, st^m_{s',m'} \ge et^m_{s,m} \end{aligned} \qquad \text{(9a)}$$

where SUU is the set containing the *physical* units u'' which have single-use constraints active; for our example $SUU = \{U1\}$. $S_{u''}$ is the index set of outlet streams for the physical unit which for our example $S_{U1} = \{S7, S8\}$. The artificial variable $asu_{u'',mu}$ is also for the physical unit and mu is the index for the number of move sets $NM_{u''}$ that require single-use constraints on the physical unit. There can also be single-use constraints on any stream and this would ensure that no overlap of moves on the same stream would occur; this is not shown. Similarly equation (9b) describes this for single-use constraints on the vessels where $SUV = \{V3\}$ and $S_{v''}$ is also for the physical vessel as $S_{V3} = \{S9, S10\}$.

$$\begin{aligned} et^m_{s,m} + et^a_{s,m} - (st^m_{s',m'} + st^a_{s',m'}) \le asu_{v'',mv} \quad \forall\ v'' \in SUV, mv = 1..NM_{v''}, s, s' \in S_{v''}, \\ m = 1..NM_s, m' = 1..NM_{s'}, st^m_{s',m'} \ge et^m_{s,m} \end{aligned} \qquad \text{(9b)}$$

It should be noted that for vessels which can accumulate inventory another set of single-use constraints could be added for each vessel if required to ensure single-use especially when a fill-draw-delay is not specified. The fill-draw-delay constraints are similar to single-use constraints where vessels with fill-draw-delay are found in equation (10) below.

$$\begin{aligned} et^m_{s,m} + et^a_{s,m} - (st^m_{s',m'} + st^a_{s',m'}) + FDD_v \le afdd_{v,mv} \quad \forall\ v = 1..NV, mv = 1..NM_v, s \in U2V, s' \in V2U, \\ m = 1..NM_s, m' = 1..NM_s{}', st^m_{s',m'} \ge et^m_{s,m}, FDD_v \ge 0 \end{aligned} \qquad \text{(10)}$$

where $afdd_{v,mv}$ is the time artificial variable and FDD_v is the fill-draw-delay in the same time unit of measure as the start and end-times. If a vessel does not require a fill-draw-delay then FDD_v should be set to any negative value. This will allow the same constraint set-up to be used for standing-gauge restrictions where $FDD_v = 0$. Moreover, the lower up-time for any stream or unit can be generated similar to equations (9) and (10) where we only show the constraint set for streams.

$$st^m_{s,m} + st^a_{s,m} - (et^m_{s,m} + et^a_{s,m}) + LUT_s \le alut_{s,m} \quad \forall\ s = 1..NS, m = 1..NM_s, LUT_s > 0 \qquad \text{(11)}$$

where LUT_s is the lower up-time value in the time unit of measure for any given stream; note that an upper up-time or maximum run-length timing constraint could also be similarly included. The final set of time reconciliation constraints relate reconciled stream move quantities with the reconciled start and end-times for a stream move and are almost identical to equation (4).

$$\sum_{t=ST_{s,m}}^{ET_{s,m}} fr_s^L \cdot (et_t^m + et_t^a - (st_t^m + st_t^a)) - \sum_{t=ST_{s,m}}^{ET_{s,m}} (f_{s,t}^m + f_{s,t}^a) \le aafr_{s,m}^L \quad \forall \; s = 1..NS, m = 1..NM_s \tag{12a}$$

$$\sum_{t=ST_{s,m}}^{ET_{s,m}} (f_{s,t}^m + f_{s,t}^a) - \sum_{t=ST_{s,m}}^{ET_{s,m}} fr_s^U \cdot (et_t^m + et_t^a - (st_t^m + st_t^a)) \le aafr_{s,m}^U \tag{12b}$$

where $aafr_{s,m}^L$ and $aafr_{s,m}^U$ are the artificial variables to be included in J_T. These penalties are different than those in equation (4) as indicated by the extra prefixed "a". There can also be a set of constraints to force all reconciled start-times to be greater than or equal to the start-time of the reconciliation and all reconciled end-times to be less than or equal to the end-time of the horizon but these are omitted given that if there are found to be in violation this is another avenue to detect and identify defects. Although there are more timing constraints that could be added to the time reconciliation model, these are dependent on the operating rules or logic constraints dictated by the production environment. It should also be further emphasized that only re-sizing and re-timing decisions have been made without re-sequencing and re-assigning decisions. If there are larger than normal adjustments made to the start and end-times that also violate the timing constraints requiring non-zero artificial variables at the solution, then re-sequencing and re-assigning decisions maybe necessary. Yet, these would have to be made as *what-if* type decisions entered manually given that an automated algorithm would be require a search technique such as branch-and-bound. Refer to Kelly (2000) for an example of how to setup of a branch-and-bound to automatically search for measurements to be deleted in mono-period data reconciliation. From the perspective of previous work on reconciliation related to scheduling, there is a recent article by Puming and Gang (2002) who stress the need to include other related constraints into the problem in order to increase the overall redundancy of the system and hence the overall *power* of detecting and isolating production defects or anomalies. On the other hand, their work required the formulation of a non-linear problem whereas in this study we have linear constraints with both formulations requiring of course a quadratic objective function.

Illustrative Example

Taking the previously described illustrative example, it is appropriate to detail how to go about using the quantity and time reconciliation QP's to diagnose defects in the past day's production for instance. Before we proceed, it is first important to highlight that when the two reconciliation models are configured for the example and the nominal or noise-free data presented in the tables is used both J_Q and J_T equal zero after solving with the weighting parameters specified in Table 7. The QP solver is *Xpress-QP* from Dash Optimization Inc. using *Xpress-Mosel* as the modeling language (Gueret *et. al.* (2002)). For the rest of the example demonstration we also deliberately exclude noise and instrumentation flaws from the data in order not to confound our presentation. If the defects we inject into the problem cannot be observed without noise then they will never be observable with noise where the converse is not true. In real situations where there will be a much smaller signal-to-noise ratio than in our example, the production analyst must always use their best judgment and include any supplementary information to increase the level of redundancy and to decrease the level of confounding.

Table 7. Weighting parameters for both the quantity and time reconciliations (inverse of the variance).

w_s	16.0 = (1/(0.25 m3))^2
w_v	16.0 = (1/(0.25 m3))^2
w_Q	10,000
$w_{st,s}$	16.0 = (1/(0.25 h))^2
$w_{et,s}$	16.0 = (1/(0.25 h))^2
w_T	10,000

Mis-Specified Stream Move

The first defect insertion is a mis-specified stream move. In Figure 2 let’s assume that the second move on stream S2 did indeed actually happen but it was mistakenly recorded as occurring on stream S1 at the same time with the same quantity as S2. This is a plausible event given that operators are sometimes too busy to properly record production activities especially when a unit operation is being shut-down at the same time (see streams S3, S4 and S7). When the quantity reconciliation is run (taking 0.2 seconds using an interior-point QP algorithm on a Pentium III 750 GHz laptop) J_Q = 3,319.2 which is a large number compared to a rudimentary estimate of the *global* maximum-power statistic threshold value (Crowe (1988)) for the objective function of 22 measured flows + 56 measured inventories = 78. The threshold or critical objective function value, assuming a normal distribution, is actually related to the Hotelling statistic however it is well-known that an approximate guess is directly proportional to its degrees-of-freedom estimate. Thus at least one defect is detected. Table 8 shows the largest adjustment to the time-period stream flows and the inventories which provides a reasonable location for the injected aberrant.

Table 8. Largest flow and inventory adjustments and artificial variables for a mis-specified defect.

$f^a_{S1,12}$	-2.0143 m3
$f^a_{S3,11}$	3.0050 m3
$f^a_{S4,11}$	-3.0114 m3
$h^a_{V1,11}$	-4.4101 m3
$h^a_{V2,11}$	5.0832 m3
$afr^L_{S1,2}$	0.0139

$afr^{L}_{S4,1}$	0.0106
$afr^{U}_{S3,1}$	0.0053

As mentioned, it is possible to compute individual statistics and threshold values for each flow and inventory adjustments yet for this example with the same weighting parameters the relative magnitudes of the adjustments are themselves comparable and revealing. As can be seen in Table 8, the mis-specification of a stream move on S1 in time-period 12 shows up as an adjustment in the S1 flow because the measured inventory in V1 is not consistent. Both the stream flow adjustments to S3 and S4 at time-period 11 are also large given the inconsistency in both V1 and V2 measured inventories if that S1 move were to have actually occurred. The artificial variables for the lower and upper flowrate bounds are also significantly non-zero implying that the reconciled flows violate the lower rate bounds on streams S1 and S4 and the upper bound on S3. When the time reconciliation is run (taking 0.3 seconds of computation time) J_T = 1.411 compared to an estimated threshold value of 24 measured start and end-times indicating no defects. Table 9 shows the largest adjustments and time artificials which even though there is no statistical defect significance, the location of the largest adjustments is consistent with the injected abnormality.

Table 9. Largest start and end-time adjustments and artificial variables for a mis-specified defect.

$st^{a}_{S1,2}$	0.0017 h
$et^{a}_{S1,2}$	-0.0017 h
$aafr^{L}_{S4,1}$	0.0106
$aafr^{U}_{S3,1}$	0.0053

Mis-Logged Stream Move

A mis-logged defect is injected by ignoring the stream move S4 at time-period 11. In reality the stream move was executed by operations except that it wasn't logged explicitly into the logging system. Again this is a probable example of a gross error in light of the fact that stream moves S3 and S7 were properly logged and an oversight was made for stream S4. After running the quantity reconciliation J_Q = 49,996.3 indicating at least one defect. Due to the very large value, an artificial is most likely non-zero as can be observed in Table 10.

Table 10. Largest flow and inventory adjustments and artificial variables for a mis-logged defect.

$f^{a}_{S3,11}$	4.5031 m3
$f^{a}_{S7,11}$	-4.4969 m3
$h^{a}_{V1,11}$	-3.0318 m3
$h^{a}_{V2,11}$	6.1520 m3

$h^{a}_{V3,11}$	-4.4969 m3
$afr^{L}_{S7,1}$	1.4970
$afr^{U}_{S3,1}$	1.5031

Without going into too much detail on the analysis of the adjustments and artificials it is very apparent that the mis-logged move is correctly isolated especially given the penalty errors for # and S7 flowrastes. Running the time reconciliation given the reconciled stream move quantities from the quantity reconciliation we also get a very large objective function of J_T = 44,471.6. As is evident from Table 11, the same large penalties on the flowrates for stream S3 and S7 exist which makes sense given that the move on stream S4 actually occurred but was failed to be recorded leaving the other two moves to account for the inventory changes in the vessels.

Table 11. Largest start and end-time adjustments and artificial variables for a mis-logged defect.

$st^{a}_{S3,1}$	0.0179 h
$st^{a}_{S7,1}$	0.0179 h
$et^{a}_{S3,1}$	-0.0179 h
$et^{a}_{S7,1}$	-0.0179 h
$aafr^{L}_{S7,1}$	1.3174
$aafr^{U}_{S3,1}$	1.6467

The very large objective function values for mis-logged moves are not unexpected. Kelly (1999) describes a three-phase procedure to help detect and identity model commissioning errors, missing or unlogged movements and bad instrumentation given production data from large-scale plant-wide flowsheets. He demonstrates using a real oil-refinery example similar large inflation of the objective function (J_Q) whenever mis-logged moves exist in the data.

Mis-Operated Stream Moves

The last production anomaly to check for are mis-operation errors or mistakes which are more relevant to the time reconciliation. Here we include two of the defects simultaneously and they are to operate vessel V2 in standing-gauge and to terminate the operation in unit U12 after 4 hours instead of 5 hours thus incurring an upper up-time violation. In this example the stream moves are recorded identical to the moves found in Table 3 and Figure 2 yielding as expected a quantity reconciliation objective function of J_Q = 0.0 because all of the flows and inventories reading are consistent with one another. However, we compute a J_T = 27.4 indicating that the operators correctly recorded all of the stream move activities for the day in question except that they mistakenly operated vessel V2 as a running-gauge tank and continued the unit operation in U12 greater than the maximum run-length of 4 hours where Table 12 shows the results. It should also be observed that the magnitude of the adjustments are very close to their individual 95% confidence interval of approximately 2 * 0.25 = 0.5 hours. This implies that the marginally acceptable global or overall test on the objective function J_T is consistent with the fact that the individual adjustments themselves are also marginally acceptable when compared to their threshold value or confidence range; a more accurate individual threshold value is to use the Student t distribution (Crowe (1988)).

Table 12. Largest start and end-time adjustments for a mis-operated defect.

$st^{a}_{S5,1}$	0.5705 h
$st^{a}_{S6,1}$	0.5705 h
$st^{a}_{S8,1}$	0.5703 h
$et^{a}_{S2,1}$	-0.4288 h
$et^{a}_{S5,1}$	-0.4274 h
$et^{a}_{S6,1}$	-0.4274 h
$et^{a}_{S8,1}$	-0.4277 h

For this case we observe that the correct start and end-time variables are perturbed from their measured values when the operating constraints in the time reconciliation problem must be respected. The end-time of the first stream move on S2 is in violation due to the standing-gauge restriction and is pushed back in time by 0.4288 hours in order to avoid the standing-gauge violation. The other three moves had their start-times pushed forward in time so as to satisfy standing-gauge operation. The end-times of the same three moves we also pushed back in time to respect the 4 hour upper up-time.

Conclusions

Presented in this article are the fine points to develop and implement a quantity and time (or logistics) reconciliation procedure solved using two quadratic programs. This procedure can potentially establish reasons for significant defects uncovered during the operation of production equipment in the process industries allowing possibly increases in throughput and decreases in overall manufacturing costs just to name a few. The methodology is designed around the quantity and logic variables and constraints of any production facility and is intended to reconcile the logistics of the manufacturing in the production-chain. A small but representative illustrative example was chronicled and it is shown that the procedure is effective at helping to diagnosis production faults such as mis-specified movements, mis-logged movements and mis-operated movements. Other defects such as mal-functioning flow meters and inventory gauges can also be readily detected by this procedure and is well-studied in the literature on data reconciliation (Kelly (2000)).

References

Albequerque, J.S. and Biegler, L.T., “Data reconciliation and gross-error detection for dynamic systems”, *AIChE Journal*, **42**, 2841-2856, (1996).

Bagajewicz, M. J. and Jiang, Q., “Integral approach to plant linear dynamic reconciliation”, *AIChE Journal*, **43**, 2546-2558, (1997).

Binder, T., Blank, L., Dahmen, W., Marquardt, W., “Regularization of dynamic data reconciliation problems by projection”, *IFAC Symposium*, ADCHEM 2000, Pisa, Italy, June, (2000).

Crama, Y., Pochet, Y. and Wera, Y., “A discussion of production planning approaches in the process industry”, Working Paper, GEMME 0102, Université de Liège, (2001).

Crowe, C.M., “Recursive identification of gross errors in linear data reconciliation”, *AIChE Journal*, **34**, 541-550, (1988).

Crowe, C.M., “Data reconciliation - progress and challenges”, *J. Proc. Cont.*, **6**, 89-98, (1996).

Fillon, M., Meyer, M., Pingaud, H. and Joulia, X., “Data reconciliation based on elemental balances applied to batch experiments”, *Computers and chem. Engng.*, **19**, Suppl., S293-S298, (1995).

Graves, S. C., “A review of production scheduling”, *Operations Research*, **29**, 646-675, (1981).

Gueret, C., Prins, C., Sevaux and Heipcke S. (revisor and translator), *Applications of Optimization with Xpress-mp*, Dash Optimization, Blisworh, Northan, UK., (2002).

Harikumar, P. and Narasimhan, S. “A method to incorporate bounds in data reconciliation and gross error detection – Parts I and II”, *Computers and chem. Engng.*, **17**, 1115-1128, (1993).

Kelly, J.D., “Practical issues in the mass reconciliation of large plant-wide flowsheets”, *AIChE Spring Meeting*, Houston, Texas, March, (1999).

Kelly, J.D., “The necessity of data reconciliation: some practical issues”, *NPRA Computer Conference*, Chicago, Illinois, November (2000).

Kelly, J.D., “On the formulation of petroleum and petrochemical scheduling optimization models”, submitted to *AIChE Journal*, (2002).

Kelly, J.D. and Mann, J.L., “Crude-oil blend scheduling optimization: an application with multi-million dollar benefits: part II”, *Hydrocarbon Processing*, July, (2003).

Maravelias, C.T. and Grossmann, I.E., “A new general continuous-time state task network formulation for the short-term scheduling of multiproduct batch plants”, submitted for publication, (2002).

Pinedo, M., *Scheduling: theory, algorithms and systems*, Prentice Hall, New Jersey, (1995).

Puming, Z. and Gang, R. “Steady-state bilinear data reconciliation dealing with scheduling”, *IFAC*, 15th Triennial World Congress, Barcelona, Spain, (2002).