A WAVELET ENHANCED INTEGRAL APPROACH TO LINEAR DYNAMIC DATA RECONCILIATION

Hsiao-Ping Huang[1], Kuo-Yuan Luo

Department of Chemical Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.

Abstract: An integral approach for dynamic data reconciliation that combines a direct numerical integration via Simpson's rule and data smoothing via discrete wavelet decomposition is presented. By simple numerical integration, the differential-algebraic equations governing the material balances are transformed into algebraic constraints to formulate the reconciliation problem. The frequency responses and the frequency contents of the measured variables are considered to determine the cut-off frequencies for data smoothing. Repetitious solutions for reconciliation using a moving data window are then used to generate the dynamic reconciled data for gross error detection. Compare with the other methods such as the Kalman filter and another sophisticated integration approach, this proposed method is simpler and has better results.*Copyright © 2005 IFAC*

Keywords: data processing, integral equation formulation, wavelet filtering

1. INTRODUCTION

Through the use of data reconciliation techniques, the corruption of process variables due to measurement noises can be reduced so as to provide correct process data and information which are useful for improving the understanding of the process and the control performance etc. Formally, data reconciliation can be defined as the estimation of variables from their measurement data to reduce measurement error through the use of the temporal and functional redundancies. Mathematically, it can be formulated as an optimal estimation problem aimed to a constrained least-squares or maximum likelihood objective function. Reviews of these methods and strategies have been given in some published books (Mah, 1990; Madron, 1992;

Romagnoli and Sanchez, 2000; Narasimhan and Jordache, 2000, etc.). In contrast to the reconciliation of data for process under steady-state, dynamic data reconciliation is more desirable, as most of the plants in running are dynamic in nature and are seldom truly under steady-state.

Methods using Kalman filter estimation, nonlinear programming, and integral approach etc. have been adopted to solve the above mentioned dynamic reconciliation problem. Among which, the Kalman filter estimation techniques are most often mentioned. Kalman filtering techniques can be adapted to take advantage of spatial and temporal redundancy (Stanley and Mah, 1977). The need for an accurate model in the formulation and in the implementation for a Kalman filter is a major bottleneck. The modelling error occurs inevitably in the procedures of linearizing nonlinear processes. Because of these methods are highly dependent on the accuracy of model, poor performances are inevitable when the process model encounter errors.

[1] Correspondence concerning this article should be addressed to H. P. Huang:
Tel: +886-2-23638999;
Fax: +886-2-23623935,
E-mail: huanghpc@ntu.edu.tw

Dynamic data reconciliation by the use of nonlinear programming (NLP) techniques is another approach. The burden and the sophistications of the computation are the main concerns of this kind of methods.

Alternatively, integral approach (Bagajewicz and Jiang; 1997) provides another choice for solving the problem. By the use of polynomial approximation, smooth results compared with other dynamic reconciliation methods can be obtained. But, it may be difficult to separate successfully the signal from noise. As a result, some significant part of the signals may be removed along with the noises, if the order of approximation is not taken sufficiently. In the other way, if the order of approximation polynomial is taken higher, a large number of computations and some huge matrices in the formulation due to a large number of parameters will be encountered. As a result, the order of the polynomial is essential and needs to be determined before the data smoothing can be performed. In fact, polynomial approximation may not fit those data that have special dynamic trends, and data filtering may be preferred.

In smoothing the dynamic data, the frequency contents of the data are important. The frequency content of any measured signal in a dynamic process is related to the dynamic lag (e.g. the time constant) of the process. As for noise filtering, the wavelet decomposition provides an excellent tool without entangling the signal with noises while filtering. A wavelet-based regulation of dynamic data reconciliation method had also been proposed (Chen et. al., 2002). It is accomplished by wavelet regulation which determines the optimal scale of the wavelet decomposition level. Then, the differential equations are converted into algebraic equations and the dynamic reconciliation problem is solved by SQP method. But, the sophistication of computation is the main difficulty of this method. Thus, filtering by wavelet-based method based on simpler way will be preferred to the smoothing by polynomial approximation, and will be used in this study. In order to facilitate the integration approach and the data smoothing for dynamic data reconciliation, an alternative approach that combines direct numerical integration and the wavelet-based filtering of dynamic data are proposed. In this proposed approach, the dynamic data are filtered using a discrete wavelet decomposition technique before computing for reconciliation. The filtering cut-off frequency is determined by the dynamics of the process. By integrating the filtered data, the differential-algebraic equations which govern the material balances are transformed into algebraic constraints. And, the reconciliation problem is solved just like the one for steady-state.

2. WAVELET FILTERING

In theory, the wavelet filtering technique is useful in pre-treating measured signals. The wavelet filtering of noisy data is accomplished by discrete wavelet transformation (i.e. DWT). A time-scale (frequency) representation of a signal is obtained using digital filtering techniques and with the combination of down-sampling operations, the time-scale representation is obtained in different frequency bands. De-noising is accomplished by reconstructing the time-scale representation with high frequency coefficients omitted according to the desired requirement. As an example, consider the level of tank no.4 in the example, the time domain reconstructions of the signal are illustrated in Fig. 1. The left column (called "Approximate") is the remained de-noised signal with different extent of the high frequency having been removed, and, the right column (called "Detail") is the corresponding removed high frequency signals at different bands. The original signal is reconstructed by combining A_1 and D_1, and A_1 is reconstructed by combining A_2 and D_2, and so on so that A_i is reconstructed by combining A_{i+1} and D_{i+1}. In dynamic systems, the original measured signals consist mainly of low frequency components and some high frequency measurement noises. As a result, the low frequency components are the main constituents of this measured signal. In fact, the energy of a signal is designated by its absolute value. Consequently, the absolute value of D_i would be lower compared with that of A_i , if the decomposition level is proper. Whenever, if the mean of the absolute value of D_i has an obvious increase at certain decomposition level, as illustrated as the hollow points in Fig. 2, the decomposition level in correspondence will not be appropriate. By this reasoning, the desired filtering extent can be determined and will be given in the text follows.
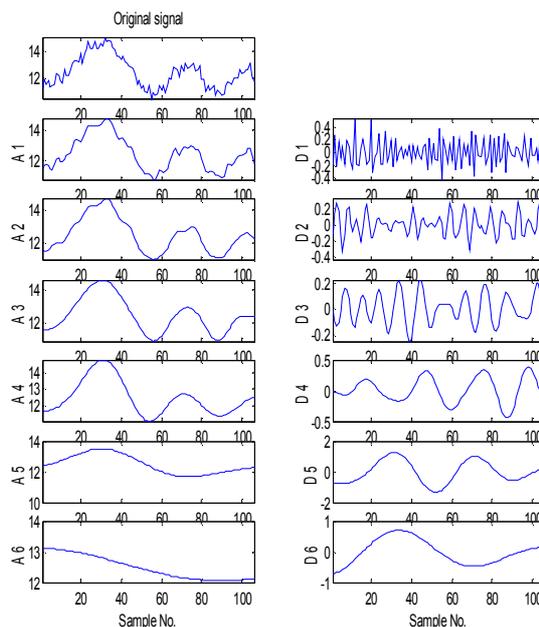

Fig. 1. DWT reconstruction: Left column---
'Approximate'; Right column --- 'Detail'
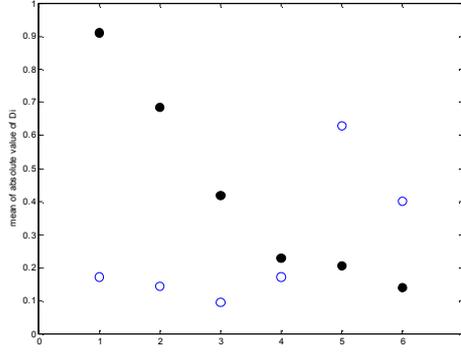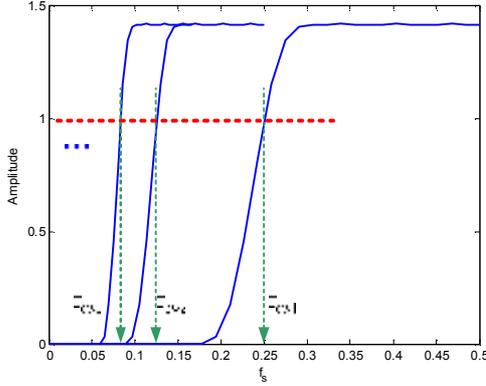
Fig. 2. Mean of the absolute value of D_i

Fig. 3. Band-pass spectrum of wavelet function

By Fourier transformation analysis of a wavelet function, the decomposition of Di appears as a high-pass digital filter. After the down-sampling operation, the wavelet function filter has a band-pass spectrum as shown in Fig. 3. The cut-off frequency, i.e. f_{cw}, of each decomposition level can thus be defined as the frequency where its magnitude equals 0.707 of the pass-band magnitude. In Fig. 3, the cut-off frequencies of the high-pass filter at different levels of wavelet transformation are designated by f_{cw1}, f_{cw2}, f_{cw3}, …, etc., and f_s stands for the sampling frequency. The exact value of f_{cw} is closely related to the sampling frequency f_s by the relations listed in Table 1. In the other words, we adopt DWT to filter the measurement signals up to certain DWT level to de-noise the measurement signals. In the following, determination of the proper DWT level to filter the measurement signals will be illustrated.

Table 1 Cut-off frequencies of different wavelet decomposition level

DWT level	Cut-off frequency (f_{cw})
1	$0.25*f_s$
2	$0.125*f_s$
3	$0.0625*f_s$
⋮	⋮
n	$0.25*f_s*(0.5)^{n-1}$

The measurement signals is a dynamic process are classified into two types.

Type 1: The first type signals are contained in the streams which are independent of the downstream processes.

Type 2: The second type is the signals contained in some down streams from some dynamic units in the process.

For type 1 measurements, the cut-off frequency corresponds to each wavelet decomposition level is defined at the frequency of suddenly increase of the mean value of D_i as explained. As an example, in Fig. 2, the decomposed D_i from the forth tank level has an obvious increasing at DWT level 4, thus, the third level of the DWT will be taken.

For type 2 measurements, the wavelet decomposition level is determined by the combination of dynamic characteristic and the frequency composition of the input signals. It is known that the dynamic behavior of a process, especially the inventory one, comes mainly from its time constants (designated as τ). The frequency response of a dynamic system is usually represented by a Bode's plot that illustrates how the process behaves like a low-pass filter that filters out the high frequency noises. Moreover, the corner frequency of each low-pass filter equals to $1/τ$. Thus, the significant frequency contain of a measured signal from a process depends on the bandwidth of the low-pass filter (i.e. process) and the input signals to the filter. In practice, the signals in the processes are mostly dominated by low frequent signals. By the Fourier analysis of a signal we can obtain its spectrum over the frequency range of the signals as shown in Fig. 4. Owing to normalizing by the total energy (i.e. the area under the curve) the amplitudes are ranged from 0 to 1.

How the type 2 signals being affected by process can be imaged with multiplying the Fourier analysis results by the Bode' plot at each frequency point. After the operation, it results the curve in Fig. 5. Due to the characteristic of the signal and the process, the amplitudes of high frequency region are much smaller than those of low frequency region and, thus, signals in the high frequency region are viewed as noises and must be filtered out. The cut-ff frequency is determined at the frequency where its amplitude is about certain ratio of the highest amplitude, which is set to 1%~5% in the examples. In Fig. 5, amp_1 is the highest amplitude and amp_2 is about 1~5% of amp_1. Then f_c is the determined cut-off frequency and the DWT decomposition level is also determined according to Table 1.
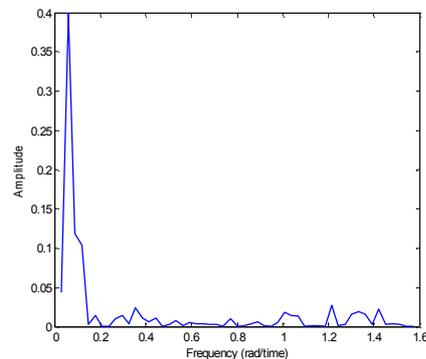

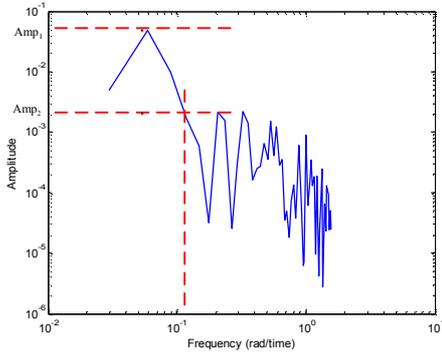
Fig. 4. Fourier analysis of a signal

Fig. 5. Frequency response through the process

3. INTEGRAL-APPROACH RECONCILIATION

Consider that a linear dynamic system be represented by the following differential-algebraic equations (DAE):

$$\frac{dh}{dt} = Af \quad (6)$$

$$Cf = 0 \quad (7)$$

In Eq. (6) and Eq. (7), h is the vector of state variables, f is the vector of non-state variables, A and C are constant matrices from the algebraic part of the DAE. By integrating Eq. (6) and Eq. (7) between some time t_1 and t_2, we can get the following algebraic Equations (i.e. Eq. (8) and Eq.(9)).

$$\int_{t_0}^{t_n} \frac{dh}{dt} dt = A \int_{t_0}^{t_n} f dt \quad (8)$$

$$C \int_{t_0}^{t_n} f dt = 0 \quad (9)$$

Notice that, in many reconciliations based on mass or energy balances, the variables (i.e. the vector f) in the right hand side of the Eq. (6) are independent of the state variables. In such cases, formulation of Kalman filter estimation will encounter almost the same difficulties as those encountered in the problems of controlling an integrating process. On the other hand, if the variable vector f is formulated in terms of the state variables based on some physical and chemical laws, the problem solution will encounter difficulties of robustness to the modeling error. To avoid these difficulties, the data is smoothed by applying the DWT first, and the filtered data is used to formulate the data reconciliation as follows.

Let $Z_1 = \int_{t_0}^{t_n} \frac{dh}{dt} dt$, $Z_2 = \int_{t_0}^{t_n} f \, dt$, and we can rearrange the integrations and get the following matrix form in Eq. (10).

$$\begin{bmatrix} A & -I \\ C & 0 \end{bmatrix} \begin{bmatrix} Z_2 \\ Z_1 \end{bmatrix} = 0 \quad (10)$$

It shows that the result is an algebraic constraint. In the following, the reconciliation procedure is to incorporate the integrating part by Simpson's rule for numerical integration.

The Simpson's n+1 (n is even) points rule is shown in Eq. (11).

$$\int_{t_0}^{t_n} f(t)dt = \frac{s}{3}\left(f_0 + 4f_1 + 2f_2 + 4f_3 + \ldots + 4f_{n-1} + f_n\right), \ s = \frac{t_n - t_0}{n} \quad (11)$$

Define new variable H and F which represent the collections of all measurements of all instruments during the integrating time interval t_0 to t_n. Assuming there are k variables (k instruments) in h and m variables (m instruments) in f and then H and F (shown in Eq. (12)) will contain $k*(n+1)$ and $m*(n+1)$ values respectively.

$$H = \begin{bmatrix} h_{1,t_0} \\ \vdots \\ h_{1,t_n} \\ \vdots \\ h_{k,t_0} \\ \vdots \\ h_{k,t_n} \end{bmatrix}, \quad F = \begin{bmatrix} f_{1,t_0} \\ \vdots \\ f_{1,t_n} \\ \vdots \\ f_{m,t_0} \\ \vdots \\ f_{m,t_n} \end{bmatrix} \quad (12)$$

The integration of Eq. (8) and Eq. (9) can be represented by H and F as represented easily in Eq. (13) and Eq. (14).

$$Z_1 = Q_1 * F \quad (13)$$

$$Z_2 = Q_2 * H \quad (14)$$

Q_1 and Q_2 are matrices shown as Eq. (15) and Eq. (16). In the two equations 's' is the sampling time interval.

$$Q_1 = \frac{s}{3} \begin{bmatrix} \overbrace{1 \ 4 \ 2 \ \cdots \ 4 \ 1}^{n+1} & \cdots & 0 \\ & \ddots & \\ 0 & & 1 \ 4 \ 2 \ \cdots \ 4 \ 1 \end{bmatrix} \quad (15)$$

$$Q_2 = \begin{bmatrix} \overbrace{-1 \ 0 \ \cdots \ 0 \ 1}^{n+1} & & 0 \\ & \ddots & \\ 0 & & -1 \ 0 \ \cdots \ 0 \ 1 \end{bmatrix} \quad (16)$$

Finally, we obtain the algebraic constraint represented by H and F from the original DAE.

$$\begin{bmatrix} A & -I \\ C & 0 \end{bmatrix} \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} \begin{bmatrix} F \\ H \end{bmatrix} = 0 \quad (17)$$

With the equality constraint, the reconciliation problem can be solved. Let's recall the general formulation of the steady-state data reconciliation problem. There is usually a model to describe the measurement variables and constraints to construct the reconciliation problem. Then, the data reconciliation problem is to estimate the state variables by minimizing a least square error problem.

The solution can be obtained using the method of Lagrange multipliers and is shown in Eq. (18).

$$\hat{x} = y - \Sigma A^T (A \Sigma A^T)^{-1} Ay \quad (18)$$

Where, Σ is covariance matrix of the measurement variables, y is the measurement, A is the constraint matrix, and $\hat{x}$ is the estimation of the measurement. The reconciliation problem is solved by Eq. (18), if we have the measurement signals, y, equality constraints, i.e. Eq. (17), and covariance matrix Σ.

The wavelet filtering is executed by batch and the integral reconciliation executed by moving-window within each batch as shown in Fig. 6. The length of the batch measurement collection must be long

enough in order to have temporal redundancy for reconciliation. As the moving-windows goes on, repeated reconciled variables will be obtained at each of the time instants. The final reconciled variables are taken as the average of the repeated reconciled variables at each time instant.
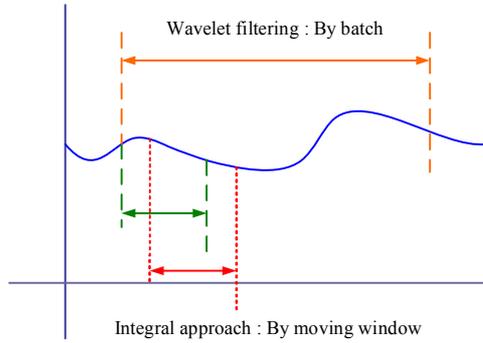

Fig. 6. Batch-like filtering and moving-window reconciliation

The length of the moving window is equal to the length of integration points. Assuming there are n+1 integration sampling points (from t_0 to t_n) and b sampling points in a batch of wavelet filtering (b>n+1).

4. EXAMPLE

A four-tank system is illustrated as an example. The DAE of this example is showed in Eq. (26). There are two main flows f_5, f_6 split into two branches apiece. The four branches, f_1, f_2, f_3, f_4, flow into four tanks respectively. Each tank has flow, q_1, q_2, q_3, q_4, out of it. The flow out of tank 3 is fed into tank 1 and the one out of tank 4 is fed into tank 2. Parameters of the process are listed in Table 2. By linearization at nominal conditions the approximated time constant values of each tank are obtained and are listed in Table 2. The type 1 variables include f_1, f_2, f_3, f_4, f_5, f_6 and the type 2 variables include h_1, h_2, h_3, h_4, q_1, q_2, q_3, q_4. According to those two groups, the DWT levels are determined by the previous two methods respectively. In this example, the two main input flows, i.e. f_5, f_6, are set to be constructed of tortuous functions with changing frequency. Assuming that there are no gross-errors and all the flows and tank levels are measured. The batch time interval of the wavelet filtering is chosen to 105 points and the moving window time interval is chosen to 22 points.

$$A_1 \frac{d h_1}{d t} = -q_1 + q_3 + f_1$$

$$A_2 \frac{d h_2}{d t} = -q_2 + q_4 + f_2 \qquad (26)$$

$$A_3 \frac{d h_3}{d t} = -q_3 + f_3$$

$$A_4 \frac{d h_4}{d t} = -q_4 + f_4$$

Table 2 Tank parameters

Symbol	State/Parameter	Value	Dimension
h_0	Nominal levels	[20.4; 20.4; 11.5; 11.5]	cm
a_i	Area of the drain	[3; 3; 2 ;2]	cm^2
A_i	Areas of the tanks	1000	cm^2
f_i	flow into the tank	[0.3; 0.3; 0.3;0.3; 0.6; 0.6]	cm^3/sec
T_i	Time constants	[68; 68; 76.5; 76.5]	sec
g	Gravitation constant	981	cm/sec^2
σ_f	Standard deviation of flow	0.015	cm^3/sec
σ_h	Standard deviation of level	0.6	cm

The DWT decomposition levels of the two different type variables are determined according to the methods mentioned previously. The mean values of the high frequency signal at different DWT level for the six variables of type 1 are calculated and the determined decomposition levels for those measurements are listed in Table 3. The cut-off frequencies of different DWT level with sampling time equal to 2 seconds in Table 5 and for type 2 variables, according to this table combined with the plot from combination of Fourier analysis of input signals and Bode's plots (time constants of each tank are listed in Table 2), the determined decomposition levels of this type variables are determined which are listed in Table 4.

Furthermore, proposed method is compared with two other dynamic reconciliation methods that are Kalman filter estimation and an integral approach proposed by Bagajewicz and Jiang in 1997. The extended Kalman filter algorithms are described in appendix. The standard deviations listed in Table 6 are calculated in order to compare the reconciliation results.

The order of the polynomial is chosen to 8, which has the best performance, in the integral approach method, and the Kalman filtering approach is under the condition of no modeling errors. In this example, our method has a better performance compared with those from Bagajewicz and Jiang's method and those from Kalman filtering approach. Bagajewicz and Jiang's method will have poorer performance as more tortuous signal existed unless the order of polynomials is creased, which is the main drawback of this method.

Table 3 Determined DWT decomposition level for type 1 variables

Variable	level	Variable	level
f_1	3	f_1	3
f_2	3	f_2	3
f_3	3	f_3	3

Table 4 Determined cut-off frequencies for type 2 variables

Variable	Cut-off frequency(Hz)	DWT level
h_1, q_1	0.0283	3
h_2, q_2	0.0283	3
h_3, q_3	0.0283	3
h_4, q_4	0.0283	3

Table 5 Cut-off frequencies of sampling time equal to 2 seconds

Wavelet decomposition level	Cut-off frequency (f_{cw}) (Hz)
1	0.125
2	0.0625
3	0.03125
4	0.015625
5	0.0078125

Table 6 Standard deviation of three methods in Ex.1

S.D.	Proposed	Integral approach	Kalman filter
q1	0.058	0.123	0.078
q2	0.053	0.132	0.080
q3	0.051	0.108	0.067
q4	0.059	0.077	0.074
f1	0.098	0.326	0.212
f2	0.086	0.440	0.249
f3	0.097	0.440	0.246
f4	0.119	0.328	0.212
f5	0.190	0.650	0.415
f6	0.173	0.877	0.485
h1	1.076	3.434	1.552
h2	1.471	4.296	1.708
h3	1.210	6.925	1.360
h4	1.273	3.826	1.450

5. CONCLUSIONS

A wavelet enhanced integral approach for the dynamic data reconciliation has been presented. Before the reconciliation computation, a filtering procedure based on DWT is applied to a batch of data for de-noising. A well known numerical integration via Simpson's rule is applied to the filtered data through a moving window to formulate the reconciliation problem. The conciliation problem is then solved sequentially to cope with the dynamic changes of the process. The results thus obtained have been compared with those using the Kalman filter estimations and those from the sophisticated integration approach. This proposed method performs better than others. On the other hand, in terms of data smoothing and reconciliation formulation, this proposed method is also much simpler than the method of Bagajewicz and Jiang (1997).

REFERENCES

G. M. Stanley, and R. S. H. Mah (1977), Eatimation of flows and temperatures in process networks. AICHE J., No. 23, pp.642.

Jose A. Romagnoli and Mabel Cristna Sanchez (2000). *Data processing and reconciliation for chemical process operations*, Academic Press, New York.

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

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

Mingfang Kong, BingZhen Chen, Xiaorong He (2002). Wavelet-based regulation of dynamic data reconciliation. *Ind. Eng. Chem. Res.*41, pp. 3405-3412.

Miguel J. Bagajewicz and Qiyou Jiang (1997). Integral approach to plant dynamic reconciliation. *AIChE Journal*, **vol. 43**, No. 10, pp. 2546-2558.

Shankar Narasimhan and Cornelius Jordache (2000). *Data reconciliation & gross error detection-An intelligent use of process data*, Gulf Publishing Company, Houston, Texas