

A Fast Method for Data Validation in RTO Technology

B. Baloochy¹, S.Shokri² and M.Ahmadi Marvast³

Abstract—Data validation is important in chemical industries. Because of random and possibly gross errors in measurements, data reconciliation is needed to minimize the measurement errors. In this article, an adaptive method is presented for dynamic and linear reconciliation of process data for real time optimization (RTO) of the process. In this method, system model parameters are estimated by the recursive least square identifier. By using these parameters, errorless process data are generated through Kalman filter method. This method is implemented using the Simulink tool box of Matlab software package. In this study, a rigorous model for the sweetening process is developed. Data generated by a sweetening Process simulation is used for evaluation of adaptive data reconciliation. First, data is artificially contaminated to errors (white noise) in Simulink environment and then filtered by the proposed method. Comparison of the outlet data for the actual and filtered process data shows great improvement and the effectiveness of the adaptive method is demonstrated.

Index Terms—Real Time Optimization, Data Reconciliation, adaptive, sweetening process

I. INTRODUCTION

Real time optimization (RTO) deals with methods of maximizing an economic objective related to the operation of a continuous process. Most RTO systems are based on non-linear steady state models of the process system, combined with data reconciliation or parameter estimation to update key-parameters [1]. The basic element of an RTO system is shown in Fig. 1.

Data validation is one of the most important blocks of an RTO system that help to estimate measured and unmeasured process variables. This block contains data reconciliation and gross error detection.

The problem of estimating the variables involved in a chemical process (data reconciliation), subject to linear balance equations, has been considered by several authors [2]-[6]. Basic issues are whether the estimation of an observed value can be improved by using the other measurements (redundancy), an unobserved value is estimable from the observed ones (determinability) and whether an observed value is a gross error. There are many

research works concerning data reconciliation in recent years [7]-[12].

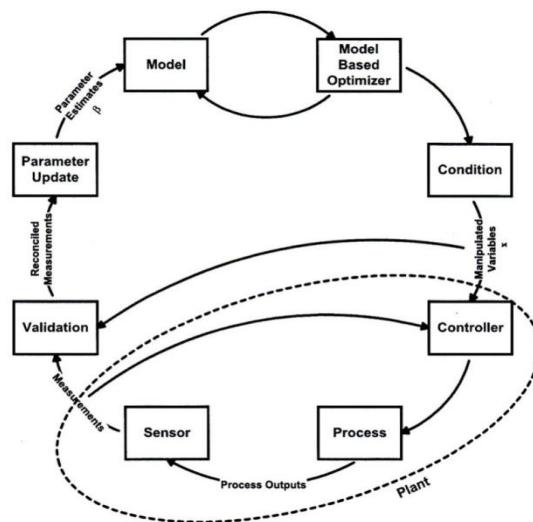


Fig. 1 Real time optimization loop

Data reconciliation is an optimization method used to eliminate random errors from measured variables. Also this method can be used to estimate unmeasured variables and model parameters. Prior to data reconciliation, it is necessary to eliminate gross errors from measured variables [1].

In data reconciliation, model equations are constraints of optimization. Based on the type of constraints, data reconciliation can be divided into two ways. According to linearity and nonlinearity of model equations, data reconciliation methods can be categorized in linear and nonlinear ways. Furthermore, based on applying time variable in model equations, data reconciliation techniques can be divided into dynamic and steady state categories [13].

Adaptive data reconciliation is a linear and dynamic method based on the above categories. The advantage of this method over other data reconciliation methods is independency of the model to the process. This method estimates its own model. Thus, unlike other methods, it is independent of the process and can be applied on any process. This technique only needs noisy values of input and output variables of the process.

II. FORMULATION PROBLEM OF DATA RECONCILIATION

Data reconciliation methods are used to improve accuracy of measured variables. What is meant by accuracy of a variable is the absolute difference between actual and measured value of variable. As stated above, data

¹Behnam.Baloochi is with Research Institute of Petroleum Industries, Tehran, Iran; (Corresponding author to provide phone: +98 21- 48253285; fax: +98 21- 44 73 97 13; e-mail: baloochyb@ripi.ir).

²Saeid.Shokri is with Research Institute of Petroleum Industries, Tehran, Iran; (e-mail: shokris@ripi.ir).

³Mahdi.Ahmadi Marvast is with Research Institute of Petroleum Industries, Tehran, Iran; (e-mail: ahmadim@ripi.ir).

reconciliation is an optimization method with objective function and constraint equations shown as follows:

Objective Function:

$$\text{Min}(Y - X)^T \Sigma^{-1} (Y - X)$$

Constraints:

$$F(X, t) = 0$$

Optimization is done using X variables. In the above equations:

F: vector function of constraints

t: time variable

Σ : variance – covariance matrix

X: measured variables vector

Y: measured values vector corresponding to measured variables of X vector

III. ADAPTIVE DATA RECONCILIATION

A. Properties of Adaptive Model

The model used in two parts of data reconciliation methods must be as general as possible to be capable of supporting as possible processes. The proposed model contains the following properties:

- 1) Black Box
- 2) Dynamic
- 3) Linear
- 4) Discrete
- 5) State Space
- 6) Multiple inputs & outputs

Model equations are as follows:

$$X(k) = AX(k - 1) + BU(k - 1) \quad (3)$$

$$Y(k) = IX(k) \quad (4)$$

$$A = \begin{bmatrix} -a_{11} & \dots & -a_{n1} \\ \vdots & \ddots & \vdots \\ -a_{1n} & \dots & -a_{nn} \end{bmatrix} \quad (5)$$

$$B = \begin{bmatrix} b_{11} & \dots & b_{m1} \\ \vdots & \ddots & \vdots \\ b_{1n} & \dots & b_{mn} \end{bmatrix} \quad (6)$$

$$X(k) = \begin{bmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{bmatrix} \quad (7)$$

$$U(k) = \begin{bmatrix} u_1(k) \\ \vdots \\ u_m(k) \end{bmatrix} \quad (8)$$

$$Y(k) = \begin{bmatrix} y_1(k) \\ \vdots \\ y_n(k) \end{bmatrix} \quad (9)$$

Where X is n dimensional vector of state variables, Y is n dimensional vector of output variables, U is m dimensional vector of input variables, I is squared identity matrix with n dimensions, A and B are matrices containing model parameters with appropriate dimensions.

B. Online Identification of Model Parameters

There are various methods to identify model parameters. In this research, recursive least square identifier is used. This method has a suitable speed convergence and keeps stability in different conditions. In this method using input and output variables of the process on latter sampling time, an estimation of model parameters is provided corresponding to the process on current sampling time. In another words, matrices A and B

are determined [14]. The identifier equations are as follows:

$$\hat{\theta}_i^T = [a_{1i} \dots a_{ni} \ b_{1i} \dots b_{mi}] \quad i = 1, 2, \dots, n \quad (10)$$

$$\hat{\theta} = [\hat{\theta}_1 \ \hat{\theta}_2 \ \dots \ \hat{\theta}_n] \quad (11)$$

$$\varphi^T(k) = [-y_1(k-1) \ \dots \ -y_n(k-1) \ u_1(k-1) \ \dots \ u_m(k-1)] \quad (12)$$

$$\text{Output}^T(k) = [y_1(k) \ y_2(k) \ \dots \ y_n(k)] \quad (13)$$

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{P(k-1)\varphi(k)[\text{Output}(k) - \hat{\theta}^T(k-1)\varphi(k)]^T}{\text{Alpha} + \varphi^T(k)P(k-1)\varphi(k)} \quad (14)$$

$$P(k) = P(k-1) - \frac{P(k-1)\varphi(k)\varphi^T(k)P(k-1)}{\text{Alpha} + \varphi^T(k)P(k-1)\varphi(k)} \quad (15)$$

Where P is squared matrix with (n + m) dimensions and Alpha is a parameter. The equations (14) and (15) are solved iteratively.

C. Estimation of State and Output Variables

Kalman filter is used to estimate state variables. Kalman filter provides an estimation of noise free state variables using identified parameters of adaptive model and noisy values of input and outputs process variables and also taking variance covariance matrix as a parameter (which is a criterion of magnitude and distribution of random errors in process variables) then using these errorless values and equation (4) errorless output variables are obtained. Kalman filter equations are as follows [15]:

$$N(k) = AM(k-1)A^T + Q \quad (16)$$

$$M(k) = N(k) - N(k)[R + N(k)]^{-1}N(k) \quad (17)$$

$$K_g(k) = N(k)[R + N(k)]^{-1} \quad (18)$$

$$Z(k) = AX(k-1) + BU(k-1) \quad (19)$$

$$X(k) = Z(k) + K_g(k)[Y(k) - Z(k)] \quad (20)$$

Where K_g is gain of Kalman filter also R and Q matrices are variance-covariance of measurement error and modeling error respectively.

IV. IMPLEMENTATION OF ADAPTIVE DATA RECONCILIATION

Each part described for adaptive data reconciliation is written as Sfunction and implemented in separate blocks in Simulink environment. Sfunction is a function with specific format written in Matlab software package to create a new block with desired performance in Simulink tool box.

The input and output process data are entered into Matlab environment from HYSYS in a suitable format using COM technology. The data is read in the Simulink tool box by workface blocks and enters the tool box environment. Then the data is artificially contaminated to random errors by white noise block. The contaminated data is primarily entered parameter identification and finally Kalman filter blocks. The output data from Kalman filter block is an estimation of errorless process output data. Fig. 2 displays the layout of block communication in Simulink tool box.

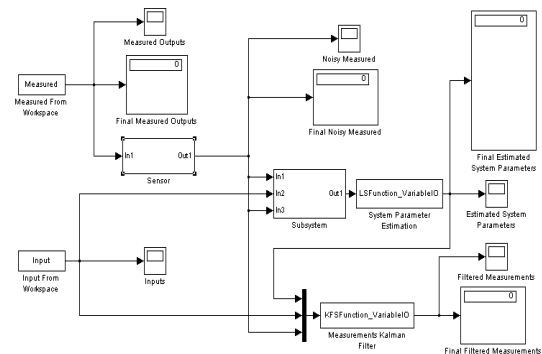


Fig. 2 Block Communications Layout

V. RESULTS AND DISCUSSION

Sweetening process is demercaptanization of petroleum fractions. In this process using caustic solutions, light mercaptans, H₂S, COS, and CS₂ are removed from the petroleum fractions and corrosive and active heavy mercaptans are converted into disulfides [16]. This process is simulated using a commercial software. Simulation results of sweetening process are used to evaluate capabilities of proposed method to eliminate random errors.

Following the entrance of input and output data from sweetening simulation to adaptive data reconciliation system, white noise signal with a power of 1000 is applied to them which are representatives of errors in measuring variables. The input data are the representatives of manipulated variables. It is assumed that the input data are deterministic and errorless.

Simulation of reconciled data is carried out discretely for 10 unit times. Sampling period equals one unit time. As simulation is done online and not real time, unit time can be seconds, minutes or any other measuring time unit.

Model Parameter identifier block consists of two arguments that are number of input and output variables. State estimation block consists of two arguments including variance-covariance matrices of modeling and measurement errors as well the above arguments. The values of these parameters are determined with respect to the measuring variable errors. As stated below:

Modeling Errors: A vector with 21 elements in length and a value of 2000 for each element

Measuring Errors: A vector with 21 elements in length and a value of 100 for each element

In Table I & II, numerical results of process variables are listed at the end of simulation for 21 process output variables.

In Table I & II, the "actual value" columns represent the primary values of process variables provided by the simulator. The "noisy value" column is corresponding values of process variables contaminated to white noise in Simulink environment. Also the "reconciled value" column contains reconciled values of noisy variables with corresponding values listed in "noisy value" column. The "difference" columns are calculated for each row as stated below:

$$\text{Noisy Difference} = |\text{Actual Value} - \text{Noisy Value}|$$

$$\text{Reconciled Difference} = |\text{Actual Value} - \text{Reconciled Value}|$$

For evaluation of reconciliation performance, mean values of the "difference" columns are calculated (Mean Absolute Error or MAE of data).

TABLE I COMPARISON OF ACTUAL AND NOISY VARIABLES

	Actual	Noisy	Difference
Controller			
1: FCV-811	20889.9	20895.0	5.1
2: TCV-811	65.0	70.2	5.1
3: LCV-811	7.4	12.6	5.1
4: TCV-812	26.1	31.2	5.1
5: FIC-101	3597.7	3602.8	5.1
6: FCV-831	25340.7	25346.0	5.3
7: TCV-831	46.0	51.1	5.1

8: LCV-831	7.5	12.6	5.1
9: TCV-832	32.6	37.7	5.1
10: FCV-851	11969.3	11974.0	4.7
11: FCV-853	7266.5	7271.6	5.1
12: LCV-851	7.9	13.0	5.1
13: TCV-851	57.9	63.0	5.1
14: FIC-105	4474.0	4479.1	5.1
15: FCV-852	109.4	114.5	5.1
16: PCV-871	144.7	149.9	5.1
17: TCV-871	54.2	59.3	5.1
18: PCV-872	390.4	395.5	5.1
19: LCV-871	55.0	60.1	5.1
20: TCV-873	51.9	57.1	5.1
21: PCV-873	945.5	950.7	5.1

Mean value of the "noisy difference" column is calculated to be 5.1.

Mean value of the "reconciled difference" column is calculated to be 3.4.

As it can be seen, reconciliation data based on new methodology has minimized the random errors imposed to variables but hasn't eliminated them completely.

TABLE II COMPARISON OF ACTUAL AND RECONCILED VARIABLES

	Actual	Reconciled	Difference
Controller			
1: FCV-811	20889.9	2080.0	1.0
2: TCV-811	65.0	67.9	2.8
3: LCV-811	7.4	10.4	3.0
4: TCV-812	26.1	28.9	2.8
5: FIC-101	3597.7	3597.4	0.3
6: FCV-831	25340.7	25340.0	0.7
7: TCV-831	46.0	48.9	2.9
8: LCV-831	7.5	10.5	3.0
9: TCV-832	32.6	35.4	2.9
10: FCV-851	11969.3	11962.0	7.3
11: FCV-853	7266.5	7260.0	6.5
12: LCV-851	7.9	10.9	3.0
13: TCV-851	57.9	62.6	4.7
14: FIC-105	4474.0	4475.6	1.6
15: FCV-852	109.4	118.7	9.3
16: PCV-871	144.7	146.9	2.2
17: TCV-871	54.2	58.8	4.6
18: PCV-872	390.4	390.3	0.1
19: LCV-871	55.0	60.0	5.0
20: TCV-873	51.9	55.3	3.3
21: PCV-873	945.5	942.0	3.5

VI. CONCLUSION

In this paper, an adaptive method is presented for dynamic and linear reconciliation of process data for real time optimization (RTO) of the process. Then the method was evaluated by standard tools. The advantage of this method over other data reconciliation methods is independency of the model to the process. This method estimates its own model. Thus, unlike other methods, it is independent of the process and can be applied on any process. This technique only needs noisy values of input and output variables of the process.

As this technique estimates model parameter online, using input and output data instantaneously, nonlinearity of the process would not cause problems in many cases. Because for any instant, linear process model is estimated around the working point of the same instant and it varies with time for any change of working point.

ACKNOWLEDGMENTS

The financial support provided by the Research and Development Center of the National Iranian Oil Company is greatly appreciated.

REFERENCES

- [1] Benqilou C., "Data Reconciliation as a Framework for Chemical Processes Optimization and Control," PhD Thesis, Universitat Politècnica de Catalunya, Spain, 2004.
- [2] Romagnoli, J. A., & Stephanopoulos, G., "On the rectification of measurement errors for complex chemical plants," *Chemical Engineering Science*, Vol. 35, pp. 1067–1081, 1980.
- [3] Crowe, C. M., Garca Campos, Y. A., & Heymak, A., "Reconciliation of process flowrates by matrix projection. Part I. Linear case," *American Institute of Chemical Engineering Journal*, Vol. 29, pp. 881–888, 1983.
- [4] van der Heijden, R. T. J. M., Heijnen, J. J., Hellinga, C., Romein, B., & Luyben, K. Ch. A. M., "Linear constraint relations in biochemical reaction systems. I. Classification of the calculability and the balanceability of conversion rates," *Biotechnology and Bioengineering*, Vol. 43, pp. 3–10, 1994.
- [5] van der Heijden, R. T. J. M., Romein, B., Heijnen, J. J., Hellinga, C., & Luyben, K. Ch. A. M., "Linear constraint relations in biochemical reaction systems. II. Diagnosis and estimation of gross errors," *Biotechnology and Bioengineering*, Vol. 43, pp. 11–20, 1994.
- [6] Romagnoli J.A., Sanchez M.C., "Data Processing and Reconciliation for Chemical Process Operations," Academic Press, 2000.
- [7] Diego Martinez Prata, Marcio Schwaab, Enrique Luis Lima, Jos é Carlos Pinto, "Simultaneous robust data reconciliation and gross error detection through particle swarm optimization for an industrial polypropylene reactor," *Chemical Engineering Science*, Vol. 65, no. 17, pp. 4943–4954, 1 September 2010.
- [8] Márcio A.F. Martins, Carolina A. Amaro, Leonardo S. Souza, Ricardo A. Kalid, Asher Kiperstok, "New objective function for data reconciliation in water balance from industrial processes," *Journal of Cleaner Production*, Vol. 18, no. 12, pp. 1184–1189, August 2010.
- [9] Flavio Manenti, Stefano Signor, Maria Grazia Grottole, Paolo Fabbri, "Adaptive Data Reconciliation Coupling C++ and PRO/II and On-line Application by the Field," *Computer Aided Chemical Engineering*, Vol. 28, pp. 373–378, 2010.
- [10] Qingfang Wu, Xavier Litrico, Alexandre M. Bayen, "Data reconciliation of an open channel flow network using modal decomposition," *Advances in Water Resources*, Vol. 32, no. 2, pp. 193–204, February 2009.
- [11] Christophe Ullrich, Georges Heyen, Carine Gerkens, "Variance of Estimates in Dynamic Data Reconciliation," *Computer Aided Chemical Engineering*, Vol. 26, pp. 357–362, 2009.
- [12] Jos é Plácido, Alexandre Almeida Campos, Daniel Ferraz Monteiro, "Data Reconciliation Practice at a Petroleum Refinery Company in Brazil," *Computer Aided Chemical Engineering*, Vol. 27, pp. 777–782, 2009.
- [13] Narasimhan S., Jordache C., "Data Reconciliation and Gross Error Detection," Gulf Publishing Company, 2000.
- [14] Ljung L., Soderstrom T., "Theory and Practice of Recursive Identification," MIT Press, 1983.
- [15] Ogata K., "Discrete-Time Control Systems," Prentice Hall, 1995.
- [16] Mazgarov A. M., Vildanov A. F., and Bazhirova N. G., "Method for the demercaptanization of petroleum distillates," US Pat. 5741415, Chevron USA Inc., 1998.