CHEMICAL ENGINEERING TRANSACTIONS, Volume 11, 2007
Edited by Sauro Pierucci

An Integrated Model-Centric Framework for Joint Parameter Estimation/Data Reconciliation of Process Systems

Daira Aragón[1,3], Pablo A. Rolandi[2] and José A. Romagnoli[1]
[1] Cain Department of Chemical Engineering, Louisiana State University. Baton Rouge, LA, U.S.A
[2]Process Systems Enterprise Limited. Bridges Studios, London, United Kingdom
[3]Departamento de Ingeniería Química, Universidad de Antioquia. Medellín, Colombia

This paper focuses on the current developments on the estimation/reconciliation modules of a novel framework for integrated decision support of process systems (IDSoPS). Built on the initial conceptual definition, a generic and versatile error-in-variables (EVM) was implemented and tested for both off-line and on-line applications using a state-of-the art modeling tool. The IDSoPS has the capability to formulate all related model-based reconciliation and estimation activities by a user not familiar with the model of the process and very little knowledge of the enabling modeling (and solution) engine. To validate these developments, the effects of random and/or systematic errors on the EVM estimation/reconciliation of an integrated chemical process are discussed.

1. Introduction

During the past two decades the computer-aided process engineering (CAPE) community has made important progresses in the development and commercialization of modeling, simulation and optimization environments (MSOEs), on the one hand, and in the establishment of standard interfaces for the communication between components of these environments, on the other hand (Braunschweig et al. 2000). One of the main features of MSOEs is the inclusion of different model-based activities (i.e. simulation, estimation, reconciliation and optimization); the corresponding mathematical problems have been subject to extensive research and have been applied widely in the past. However, one could argue that software applications targeted to process industries have not been designed to the formulation of real process engineering problems and, therefore shifting from simple to more realistic applications demands innovative and effective ways of solving old problems using current technologies and contemporary research advances.
The conceptual definition of a single and consistent model-centric framework integrated decision support of process systems (IDSoPS) aimed at facilitating the realistic formulation of related engineering problems was proposed by Rolandi and Romagnoli (2005). The framework introduces a new concept, the *Problem Definition Environment* (PDE), to provide innovative mechanisms, the *data model template* (DMT) and the *data*

model definition (DMD) for such formulation. Not only is the formulation of related phenomenological problems possible but also of entire complex chemical plants.

This paper focuses on the current developments on the estimation/reconciliation module of the framework mentioned previously. Built on the initial conceptual definition, a generic and versatile error-in-variables (EVM) was implemented for both off-line and on-line applications. To validate these developments, the effects of random and/or systematic errors on the EVM estimation/reconciliation of an integrated chemical process are discussed. It is implied through the course of this paper, that the PDE was used to formulate all model-based problems presented. This paper is organized as follows. Section 2 gives an insight of the proposed framework along with the advantages of its utilization on the problem definition. Next, an overview of the EVM data reconciliation module is presented. Then, section 3 describes the example process and presents and discusses the results obtained. Finally, section 4 draws some conclusions and the direction for future work.

2. The Framework for Integrated Decision Support of Process Systems

2.1 Framework architecture

The (framework for) IDSoPS is founded on the interaction among four main entities (see fig. 1): (i) the user, (ii) the PDE, (iii) the PME and (iv) the PMCs. The user interacts directly with the PDE to define the problem of interest, including the activity to perform and the establishment of all possible initial guesses and variable bounds. The PDE provides the mechanisms for the translation of the formulated problem into a form which is readable by the process modeling environment (PME, software tool to assist in the construction of the process model and the execution of model-based activities). Then, it delegates the solution of the problem to the corresponding PME, which coordinates the execution of the necessary process modeling components (PMCs, software components in charge of a specific function) to eventually achieve a solution.

The translation of the problem description into the *problem input files* (PIFs) readable by the PME is attained by means of two proposed new data entities: the *Data Model Template* (DMT) and the *Data Model Definition* (DMD). On the one hand, the DMT corresponds to a data structure containing all available structural and numerical information, such as the type of activity, available process variable, and statistical model, among others; furthermore, it relates the physical process variables and the mathematical variables (used for the construction of the model). On the other hand, the DMD represents an abstraction or mapping of the problem incorporating the DCS plant data in a seamless way. This structure is reusable, so that different model-based activities can be defined using the same DMD. At the current stage of the problem definition environment, both DMT and DMD are given by databases implemented in Excel workbooks files. During the *problem definition* (see fig. 1), the information contained in the DMT is presented to the user who, then, through the PDE is able to define his/her specific problem in terms of the (real) physical variables instead of the (abstract) mathematical variables, which represents an important advantage for the process engineer. During the *problem translation* (see fig. 1), the PDE creates the necessary PIFs making use of the problem specification given by the DMD. Finally, the

MSOE instatiates the problem and coordinates its execution (*problem instantiation* and *problem solution*).

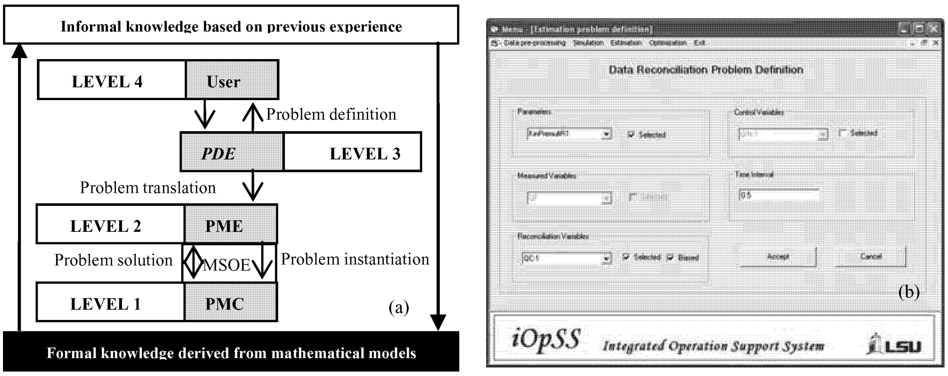


Figure 1. (a)Framework architecture and (b) Problem definition environment for data reconciliation/parameter estimation

2.1 Formulation of Error-in-variable (EVM) data reconciliation/parameter estimation problem

Data reconciliation refers to the process of obtaining better measurement estimates so that they comply with the mass and energy balances. Classical reconciliation assumes that the independent (input) process variables are error-free. However, this assumption is not valid for most of industrial processes (Kim et al., 1991). A more general approach to this problem has been studied by a number of researchers (Kim et al., 1991; Tjoa and Biegler, 1992; Romagnoli and Sanchez, 2000). In this approach, a distinction between dependent and independent variables does not exist and the measurements can be related to the errors as:

$$\tilde{z}(t) = z(t) + \varepsilon + \beta \tag{1}$$

where $\check{z}(t)$ refers to the variable measurements, $z(t)$ is the true value and ε and β correspond to random and gross errors respectively. This approach is known as error-in-variables method (EVM). In our formulation we propose the following mathematical definition for the general *estimation problem*:

$$\min_{\theta,\beta,\omega,\gamma} \varphi(\tilde{z}(t), z(t), \sigma(t))$$

subject to: (p1)

$$F(\dot{x}(t), x(t), y(t), u(t), p, \theta, \beta) = 0,\ t \in [0, t_f]$$

$$I(\dot{x}(0), x(0), y(0), u(0), p, \theta, \beta) = 0$$

$$\sigma(t) = \sigma(\tilde{z}(t), z(t), \omega, \gamma),\ t \in [0, t_f]$$

In problem (p1), $\varphi(.)$ is a generic objective function. z designates experimental observations; θ are model parameters. In our formulation these parameters can also be part of the estimation problem in the case of joint data reconciliation parameter estimation. The variables ω and γ are associated with the statistical information about the experimental observations. These parameters can also be included in the estimation problem, thus allowing the characterization of the statistical properties of the different measurements directly from plant data. *F(.)* and *I(.)* denote the set of partial differential-algebraic equations encompassing the fundamental process model and the set of initial conditions respectively. In these equations x and y denote the differential and algebraic variables respectively; $u(t)$ are the set of input variables. Additionally, $\sigma(t)$ denotes the variance model of the measurement error. The objective function can take several forms depending on the nature of the mathematical model, and on the decision variables of interest. The latter will define the name given to the general problem stated by the set of equations (p1). For instance, when only the model parameters, θ, are taken into consideration the problem will be called *parameter estimation*; if the measurement biases, β, are the only decision variables involved, the problem will now be called *gross error estimation*. A weighted least squares (WLS) objective function was implemented to solve for the measurements, (gross and random) errors and parameters (in combination or separately).

Online data reconciliation was implemented with a moving window or history horizon approach. A time horizon $H\Delta t$ is selected so that the H most recent measurements are reconciled over the time period $t-H\Delta t$; where t is the actual time and Δt is the time interval for measurement availability. The difference between the problem definition of an offline and online DDR through the PDE is the specification of the number of data in the window, H, which was set to 15 for the case study under consideration. It is worth to mention that all related model-based reconciliation and estimation activities become transparent to the user, not familiar with the model of the process and very little knowledge of the enabling modeling (and solution) engine, through the use of the proposed framework described in Section 2. In this way, the definition and redefinition of the problems to be considered as well as the variables and parameters to be used as decision variables is straightforward.

2.3 Implementation considerations

The general process modeling system, gPROMS (PSE Ltd., 2004), was selected as the state of the art modeling, simulation and optimization engine (MSOE) to solve the related model-based activities. The general EVM data reconciliation problem is not currently supported by the gPROMS language and, consequently, it was necessary to create the mechanisms for this activity. Since data reconciliation is by itself an optimization activity, this notion was used to reformulate the reconciliation problem in gPROMS. The sequential quadratic programming method incorporated in gPROMS was used to solve the EVM data reconciliation problem.

3. Case Study

3.1 Process description

The case study considered here consists of an integrated plant (two CTRS in series and fresh feed mixing with the outlet stream of the first reactor before entering the second vessel, plus a series of heat exchangers). Both a real hybrid pilot scale (real plant-soft reactions) and virtual versions of the process are available for testing, however, only the simulated results are presented here. The corresponding model for the simplified diagram has been presented elsewhere (Romagnoli and Sánchez, 2000). However, the value of the reaction parameters and input variables were taken from Bequette (1998). Additionally, the process variables were scaled using a nominal reference temperature T_r and a reference volumetric flow rate Q_r to make all variables of the same order of magnitude. Once the model was constructed in gPROMS, a simulation was run for a total of 15 hours to obtain 150 experimental values. This original dataset was modified by adding random errors to all variables so that a new data set was formed (NB). Then, measurement biases (systematic errors) were added to (two input) process variables forming a second data set (BI).

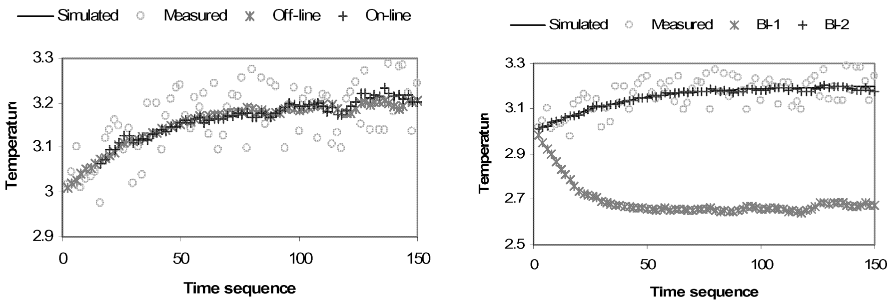


Figure 2. Comparison of reconciled trajectory for EVM-DDR for temperature in second reactor for (a) data set not biased and (b) data set with bias.

3.2 Results and discussion

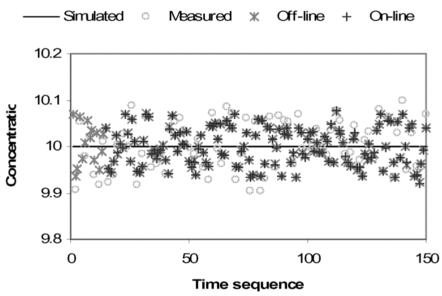
On- and off-line error-in-variables dynamic data reconciliation (EVM-DDR), and Joint parameter estimation and (EVM) data reconciliation were performed for the corresponding datasets. The reaction kinetic constant in first CSTR (CSTR1.k) and the heat transfer coefficient in second CSTR (CSTR2.UA) were considered for parameter estimation. Furthermore, measurement biases were estimated in conjunction with parameters (refer to table 1).

Table 1. Parameter and bias estimates and for joint PE/EVM-DDR

Dataset/case	Error norm	Process parameters#		Measurement biases#	
		CSTR1.k ($9702.78\ s^{-1}$)	CSTR2.UA (0.8717 kJ/s.K)	Input-1 ($1.5\ m^3/s$)	Input-2 (-0.4 K/K)
NB	658.584	9552.96	0.984	-	-
BI-1	8167.322	10500.00	0.300	-	-
BI-2	1304.910	10088.40	0.710	1.000	-0.387

#True (simulated) values in parenthesis

Table 1 shows that biases in input variables have a detrimental effect in the reconciliation/estimation results. In fact, error norm increased in approximately 94%, and parameter estimates degrades when measurement biases in input variables are not estimated (NB vs. BI-1). An improvement in error norm and in parameter estimates is observed when biases are also estimated (BI-2). In figure 2(a) the results for EVM-DDR on- and off-line for the data set with no bias (NB) are compared. It can be observed that trajectories do not differ significantly from the simulated (true) trajectory, indicating the satisfactory performance of the EVM-DDR for both off- and on-line. Figure 2(b) shows the negative effect of measurement biases over the reconciled trajectory, which does not follow the true trajectory when biases are not estimated (BI-1). This trajectory is followed closely when biases were estimated (BI-2). Figures 3(a) and (b) present the reconciled trajectory for the input concentration. The standard deviation of this variable decreases in approximately 30% once the EVM-DDR was performed. This suggests that EVM-DDR would produce more accurate estimates and, therefore, it should be adopted when input variables contain high levels of measurement noise.



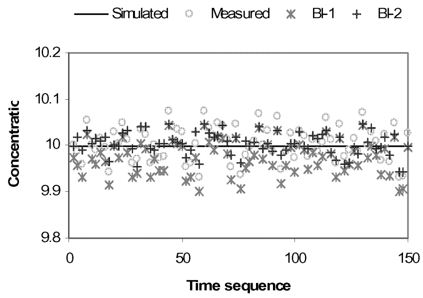


Figure 3. Comparison of reconciled trajectory for input concentration for (c) data set not biased and (d) data set with bias.

4. Conclusion

A single and consistent model-centric framework for integrated decision support of process systems (IDSoPS) has demonstrated to have the capability to formulate different model-based related problems. The error-in-variables method (EVM) was implemented within the framework for joint parameter estimation/dynamic data reconciliation (JPEDR) using the sequential approach for its solution. This method showed a good performance when measurement biases are estimated in conjunction with process parameters. A significative reduction in measurement noise of input variables was observed, which contributed to the improvement of estimates. The incorporation of robust methodologies for estimation/reconciliation activities is currently under development.

4. References

Bequette, B.W. (1998). *Process dynamics*, Prentice Hall PTR, New Jersey.

Braunschweig, B.L.; C. Pantelides; H.I. Britt and S. Sama (2000). Process Modelling: the promise of open software architectures. *Chem. Eng. Process,* **96**, pp.65-76.

Kim, I.; M.J. Liebman and T.F. Edgar (1991). A sequential error-in-variables method for nonlinear dynamic systems. *Comp. Chem. Eng.,* **15**, pp.663-670

PSE, Ltd. (2004). gPROMS advanced user guide release 2.3.

Tjoa, I. and L. Biegler (1992). Reduced successive quadratic programming strategy for errors-in-variables estimation. *Comp.Chem.Eng.,* **16**, pp.523-533.

Rolandi P.A. and J.A. Romagnoli (2005). An integrated environment for support of process operations. *AIChE annual meeting,* Cincinnati.

Romagnoli, J.A. and M.C. Sanchez (2000). "Data Processing and Reconciliation for Chemical Process Operations", Academic Press International, San Diego.