View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Repository of the Academy's Library

HUNGARIAN JOURNAL
OF INDUSTRY AND CHEMISTRY
VESZPRÉM
Vol. 41(1) pp. 65-75 (2013)

ROLE OF STEADY STATE DATA RECONCILIATION IN PROCESS MODEL DEVELOPMENT

BARBARA FARSANG✉[1], SÁNDOR NÉMETH[1], AND JÁNOS ABONYI[1]

[1]University of Pannonia, Department of Process Engineering, H-8200 Veszprém, Egyetem Street 10., HUNGARY
✉E-mail: farsangb@fmt.uni-pannon.hu

In chemical and hydrocarbon industry operational efficiency is improved by model-based solutions. Historical process data plays an important role in the identification and verification of models utilized by these tools. Since most of the used information are measured values, they are affected by errors influencing the quality of these models. Data reconciliaton aims the reduction of random errors to enhance the quality of data used for model development resulting in more reliable process simulators. This concept is applied to the development and validation of the complex process model and simulator of an industrial hydrogenation system. The results show the applicability of the proposed scheme in industrial environment.

Keywords: data reconciliation; flowsheeting simulator; validation; balance equation; industrial hydrogenation system

Introduction

In recent years, development of industrial technologies has been determined by the proliferation of computing and information technology. Today's technologies are characterized by widespread application of process engineering tools. The success of these tasks depends on the accuracy of data and applied models [1]. Thanks to the evolution of information technology, on-line and historical process data — coming from chemical process systems — are available. The collected data provide the opportunity for engineers to better understand the processes, anomalies, and malfunctions [2]. Monitoring of process variables allows us to ensure the consistent product quality. Collected data can also be used for the development and validation of process simulators. However, measurements are always affected by errors during the measurement, processing and transmission of the measured signal. Errors (gross, random, bias) in measured data affect the quality of process models and can lead to significant decrease in plant performance. Estimation of true conditions of process states is important to achieve optimal process monitoring, control and optimization.

Therefore, several methods have been developed to minimize measurement errors thereby enhancing the reliability, accuracy and precision of data. First, analogue and digital filters were used to reduce the effect of high frequency noise [3]. Gross errors were detected with various data validation methods, which include checking whether the measured data and the rate at which it is changing is within predefined operational limits. Nowadays, smart sensors are used for determine whether there is any hardware problem or the measured data are appropriate. New methods were statistical quality control tests that are applied to each measured variable separately. Although these methods improve the reliability and accuracy of the measured data, they do not ensure consistency of the data with respect to the inter-relationships between different process variables. Therefore new methods – data reconciliation and gross error detection – are developed in chemical engineering [4]. The main difference between data reconciliation and other filtering methods is that data reconciliation uses process model constraints and that the results satisfy constraints and balance equations. Data reconciliation techniques can be applied to reduce random errors of measurements, while the other techniques mentioned above do not. Gross error detection can be used for eliminate systematic errors so simultaneous data reconciliation and gross error detection have emerged as a key of online optimization [5] (see *Fig. 1*).

Data reconciliation techniques take minimal corrections of the measured variables to satisfy a set of model constraints. Based on the difference between the measured and reconciled data, the following questions can be answered. Can we consider a related set of measurements acceptable based on our previous knowledge of the system? Are the measurements consistent? If not, what can be the source of the error? Based on the available measurements and a prior knowledge, what is the most likely state of the system? Answers to these questions are important in development of technology (e.g., monitoring, optimization, simulation, control, instrument maintenance). Using this technique, we can verify the acceptability of measurements, improve the accuracy of measurements, estimate model parameters and unmeasured variables, and it can be used

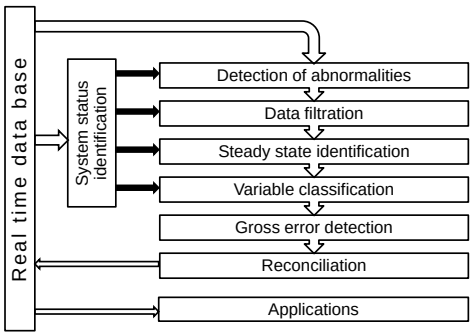

Figure 1: General methodology for on-line data correction [6]

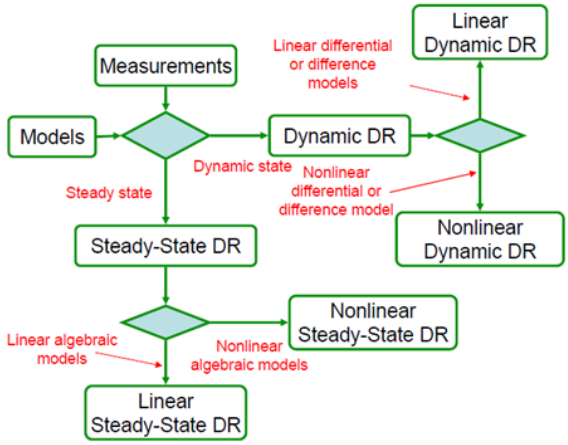

Figure 2: Types of data reconciliation [7]

for fault detection.

In order to ensure the consistency of measured process variables, models are used in data reconciliation. During data reconciliation, information is obtained from both measurements and process models. Depending on the types of the models, data reconciliation techniques can be separated into two different problems: steady state and dynamic (see *Fig. 2*). In the case of steady state data reconciliation, model constraints are algebraic equations. If we are dealing with dynamic processes, differential equations are used as model constraints. In the case of linear data reconciliation, model constraints are defined by linear equations. The simplest data reconciliation technique is the steady state linear method that can describe, e.g., the mass balance of technology. Moreover, most chemical processes have nonlinear characteristics and constraints are nonlinear equations; this problem can be called as nonlinear data reconciliation.

The goal of the presented research is to propose a methodology to support the development and validation of complex process models and simulators based on increasing the quality of measurement data used for (kinetic) parameter identification and validation. In section *Literature survey*, a historical background of steady data data reconciliation is briefly reviewed. The basic method of linear and nonlinear steady state data reconciliation is described in section *Theoretical background of steady state data reconciliation*. Data reconciliation cannot be used in process development only, but it is a useful tool in model improvement too. The principle of this application is introduced in section *Combined application of data reconciliation and flowsheeting simulator*. Our approach aims the reduction of random errors to enhance the quality of data using data reconciliation and flowsheeting simulator simultaneously. The method is described in subsection *The developed model based data reconciliation technique*. The proposed approach is illustrated based on an industrial hydrogenation system. In section *Case study*, the analyzed technology is introduced and the applicability of the method is illustrated on the basis of the mass balance of the hydrogenation system. Then, some consequences are drawn in section *Conclusion*.

Literature survey

KUHAN and DAVIDSON presented the first study about minimizing measurement errors in chemical process industry using a data reconciliation technique in 1961 [4]. They described the steady state linear data reconciliation technique and demonstrated a general solution for mass balance of technology. This seminal paper has started a new research area in process engineering. NOGITA [8] and MAH [9] improved a new method capable of detection of gross errors. Later, MAH [10] found that the base method requires a process system in steady state conditions and that the process constraints must be linear. CAO and RHINEHART confirmed that the methodology will be successful if process data come from steady state conditions [11]. In these cases, data reconciliation problems can be solved with standard numerical methods like sequential quadratic procedures [12] and quasi-Newton algorithms [13]. A multiple Gauss-Jordan elimination algorithm was published by MADRON and VEVERKA [14]. The classification method is based on matrix decomposition and permutation. The linear method was improved by CROWE (1986) who published a new technique to solve nonlinear data reconciliation using matrix projection [15]. Furthermore, BAGAJEWICZ and JIANG [16] presented an improved method that is applicable for dynamics system. The developed method from PRATA [17] is suitable for examination of dynamic nonlinear process systems.

Despite of the numerous good reviews of data reconciliation techniques, industrial application of this method started only in the 1980s. There are several complex software packages supporting data

Table 1: Industrial application of data reconciliation (newer industrial scenario and publications are added to PRATA's summary [18])

Industrial scenario	Author (year)
Absorption refrigeration systems	MARADIAGA et al. (2013) [19]
Beverage alcohol distillation plant	MEYER et al. (1993), NOUNOU and BAKSHI (1999), SCHLADT and HU (2007)
Chemical extraction plant	HOLLY et al. (1989)
Ethylene and ammonia plant	SANCHEZ et al. (1992), SANCHEZ and ROMAGNOLI (1996), PLÁCIDO and LOUREIRO (1998)
Exxon chemical process	MCBRAYER et al. (1998) , SODERSTROM et al. (2000)
Gas pipeline systems	BAGAJEWICZ and CABRERA (2003)
Gases network in an iron and steel making plant	YI and HAN (2004)
Hydrogen plant	BUSSANI et al. (1995), CHIARI (1997), SARABIA et al. (2012) [20]
Industrial coke-oven-gas purification process	FABER et al. (2006) HU and SHAO (2006)
Industrial distillation column	ISLAM (1994), WEISS et al. (1996), SANCHEZ et al. (1996), BOUROUIS et al. (1998), LI et al. (2001), BHAT and SARAF (2004), CHATTERJEE and SARAF (2004), CHEN et al. (2013) [21], KELLER et al. (2012) [22]
Industrial ETBE reactor	DOMINGUES et al. (2012) [23]
Industrial furnace	PIERUCCI et al. (1996), EKSTEEN et al. (2002)
Industrial hydrometallurgical plants for a gold extraction	DE ANDRADELIMA (2006)
Industrial polymerization reactor	VIEIRA et al. (2003), PRATA et al. (2006, 2008, 2009, 2010 [24])
Industrial synthesis gas for production of ammonia	CHRISTIANSEN et al. (1997)
Industrial utility plant	LEE et al. (1998)
Methyl-terc-butyl-ether plant	AL-ARFAJ (2006)
Mineral and metallurgical plants	VASEBI et al(2012) [25]
Nuclear power reactor	VALDETARO (2011) [26]
Refinery	PICCOLO and DOUGLAS (1996), ZHANG et al. (2001)
Sulfuric acid plant	CHEN et al. (1998), OZYURT and PIKE (2004)
Turbine cycle of a boiling water reactor	SUNDE and BERG (2003)
Vinyl acetate and ketene plants	DEMPF and LIST (1998)
Water processes	MARTINS et al. (2010) [27]

reconciliation (for instance SigmaïňĄne, DataCon, Vali, Inlibra). These tools have interfaces to information systems used to manage process data and have special model building functionalities that allow almost automatic building of plant-wide balance equations. Nowadays, data reconciliation is widely applied in various processing industries (e.g. refining, chemical industry, metals, mining, and power industry). Despite of this fact, relatively few articles deal with industrial applications (PRATA et al. prepared a summary in 2009 [18]; newer applications and papers are collected in *Table 1*).

STANLEY et al. classified variables as observable and unobservable [28]. CROWE categorized the variables from another point of view: the basis of classification is the given measured variable can be calculated from other measured variables using process models or not. The names of the two classes are redundant and non-redundant measurements [29]. ALI and NARASIMHAN applied graph theory to analyse the sensor network, classify the variables, and compute the redundancy degree of each variable. They claim that not only measured variables can be redundant. Those unmeasured variables that can be estimated in multiple ways are also termed as redundant [30].

Theoretical background of steady state data reconciliation

In this section, the basis of linear and nonlinear data reconciliation techniques are described. Different values are distinguished. Real values (without error) of variables are designated by diacritic caron (ˇ), measured values by tilde (˜), while estimated values by circumflex (ˆ).

In general, optimal estimates for process variables by data reconciliation are solutions to a constrained least-squares or maximum likelihood objective function, where measurement errors are minimized with process model constraints. The steady state data reconciliation problem can be formulated as an optimization problem by minimizing

$$J(\check{y},\check{z}) = (\tilde{\mathbf{y}} - \check{\mathbf{y}})^T \mathbf{V}_{\tilde{\mathbf{d}}}^{-1} (\tilde{\mathbf{y}} - \check{\mathbf{y}}) \tag{1}$$

subject to

$$\mathbf{f}(\check{y},\check{z}) = 0 \tag{2}$$

and

$$\mathbf{g}(\check{y},\check{z}) \geq 0. \tag{3}$$

The solution of the optimization problem is performed with the following simplifying assumptions

- The measurement error is independent from the balance variables.
- The expected value of measurement error is zero.
- The measurement error is in a normal (Gauss) distribution.
- Errors of measurements are independent from each other (diagonal covariance matrix).
- The covariance matrix is positive definite, so its inverse matrix exists.

In the following two subsections, solution methods are presented in the cases of linear and nonlinear models.

Linear data reconciliation

In the linear case, when all variables are measured, *Eq. (1)* means minimizing

$$J(\check{y},\check{z}) = (\tilde{\mathbf{y}} - \check{\mathbf{y}})^T \mathbf{V}_{\tilde{\mathbf{d}}}^{-1} (\tilde{\mathbf{y}} - \check{\mathbf{y}})$$

subject to

$$\mathbf{A}\check{\mathbf{y}} = 0. \tag{4}$$

Gauss distribution $(\tilde{\mathbf{d}} \sim N\left(0, \mathbf{V}_{\tilde{\mathbf{d}}}\right))$ is assumed, so the density function of measurement error is

$$f\left(\tilde{\mathbf{d}}\right) = C \cdot \exp\left[-0.5 \cdot \tilde{\mathbf{d}}^T \mathbf{V}_{\tilde{\mathbf{d}}}^{-1} \tilde{\mathbf{d}}\right] \tag{5}$$

from which it follows that

$$f\left(\tilde{\mathbf{y}}\right) = C \cdot \exp\left[-0.5 \cdot (\tilde{\mathbf{y}} - \check{\mathbf{y}})^T \mathbf{V}_{\tilde{\mathbf{d}}}^{-1} (\tilde{\mathbf{y}} - \check{\mathbf{y}})\right] \tag{6}$$

and that the

$$\check{\mathbf{f}} = \mathbf{A}\check{\mathbf{y}} - \mathbf{b} = 0 \tag{7}$$

criterion is satisfied in accordance with the balance equation. The essence of the most likely estimation: if $\tilde{y}$ is given and its density function is known, which $\tilde{y} \to \hat{y}$ parameter of density function will be $\mathbf{f}$ (function vector of equality model constraints).

Since the logarithm function is strictly monotonously increasing, the maximum of the logarithm of the $f(y)$ function recording only positive values is at the same place where the maximum of the argument is. Exploiting it on the above function and using the Lagrange multipliers method, the result is a vector equation system with two unknowns:

$$\mathbf{V}_{\tilde{\mathbf{d}}}^{-1}\,\hat{\mathbf{y}} - \mathbf{A}^T\hat{\lambda} = \mathbf{V}_{\tilde{\mathbf{d}}}^{-1}\tilde{\mathbf{y}} \tag{8}$$

$$\mathbf{A}\hat{\mathbf{y}} = \mathbf{b} \tag{9}$$

The solution of the vector equation system is

$$\begin{aligned} \hat{\mathbf{y}} &= \left(\mathbf{I} - \mathbf{V}_{\tilde{\mathbf{d}}}\mathbf{A}^T\left(\mathbf{A}\mathbf{V}_{\tilde{\mathbf{d}}}\mathbf{A}^T\right)^{-1}\mathbf{A}\right)\tilde{\mathbf{y}} \\ &+ \mathbf{V}_{\tilde{\mathbf{d}}}\mathbf{A}^T\left(\mathbf{A}\mathbf{V}_{\tilde{\mathbf{d}}}\mathbf{A}^T\right)^{-1}\mathbf{b} \end{aligned} \tag{10}$$

Thus, the estimation requires a matrix-vector multiplication and a vector addition. Because $\tilde{\mathbf{y}}$ is only the variable, in the other operations there are only constants [31].

In practice, not all streams and properties are measured in plant due to physical and economical reasons. The data reconciliation technique is suitable for estimating unmeasured variables. In this case, the optimization problem can be solved by the method of

Projection Matrix. In the first step, the incidence matrix is separated into two parts:

$$\mathbf{A}_y\check{\mathbf{y}} \; + \; \mathbf{A}_z\check{\mathbf{z}} = 0. \tag{11}$$

Unmeasured flows ($\check{\mathbf{z}}$) should be eliminated by pre-multiplying both sides by a projection matrix, such as $\mathbf{PA}_z\check{\mathbf{z}} = 0$. *Eq. (4)* then can be rewritten as

$$\min J(\check{y}, \check{z}) = (\tilde{\mathbf{y}} - \check{\mathbf{y}})^T \, \mathbf{V}_{\tilde{\mathbf{d}}}^{-1} \, (\tilde{\mathbf{y}} - \check{\mathbf{y}}) \tag{12}$$

subject to

$$\mathbf{PA}_y\check{\mathbf{y}} = 0. \tag{13}$$

The projection matrix is from Q-R factorization of matrix $\mathbf{A}_z$:

$$\mathbf{A}_z = \mathbf{QR} = [\mathbf{Q}_1 \; \mathbf{Q}_2] \; [\mathbf{R}_1 \; \mathbf{0}]^T \tag{14}$$

The projection matrix is the transformation of $\mathbf{Q}_2$ matrix ($\mathbf{P} = \mathbf{Q}_2^T$). If $\mathbf{A}$ matrix is replaced by matrix $\mathbf{PA}_y$ in *Eq. (10)*, then

$$\begin{aligned} \hat{\mathbf{y}} \; &= \; \left(\mathbf{I} - \mathbf{V}_{\tilde{\mathbf{d}}}(\mathbf{PA})^T \left((\mathbf{PA})\mathbf{V}_{\tilde{\mathbf{d}}}(\mathbf{PA})^T\right)^{-1} (\mathbf{PA})\right) \tilde{\mathbf{y}} \\ &+ \; \mathbf{V}_{\tilde{\mathbf{d}}}(\mathbf{PA})^T \left((\mathbf{PA})\mathbf{V}_{\tilde{\mathbf{d}}}(\mathbf{PA})^T\right)^{-1} \mathbf{b}. \end{aligned} \tag{15}$$

After reconciled values of measured variables are obtained, the next step is to estimate the unmeasured variables using the information from process models (*Eq. (11)*) as

$$\mathbf{A}_y\check{\mathbf{y}} \; = -\mathbf{A}_z\check{\mathbf{z}} \tag{16}$$

This linear equation can be solved to get estimated values of the unmeasured variables. Usually, the number of equations is greater than the number of unmeasured variables. The least-squares technique can be applied and the solution is

$$\hat{\mathbf{z}} = -(\mathbf{A}_z^T\mathbf{A}_z)^{-1}\mathbf{A}_z^T(\mathbf{A}_y\hat{\mathbf{y}}). \tag{17}$$

Nonlinear data reconciliation

The data reconciliation problem can be extended to nonlinear steady state models. It is necessary because processes in chemical industry cannot be described only by linear models; behaviour of chemical processes are often nonlinear. If we wish to simultaneously reconcile mass flow, composition, or temperature measurements, mass, component, or energy balances have to be included as constraints. Sometimes, inequality constraints have to be defined (e.g. mass/mole fraction of components have to be in the [0:1] interval). Moreover, if we take thermodynamic equilibrium relationships and complex correlations for thermodynamic and physical properties as constraints, nonlinear data reconciliation techniques have to be used.

If all variables are measured, the bilinear data reconciliation problem can be reduced to a linear problem by introducing the "measured" compound flows ('component flows' or 'energy flows'). Therefore, solution of the nonlinear data reconciliation problem is presented on the basis of the situation when measured and unmeasured variables occur.

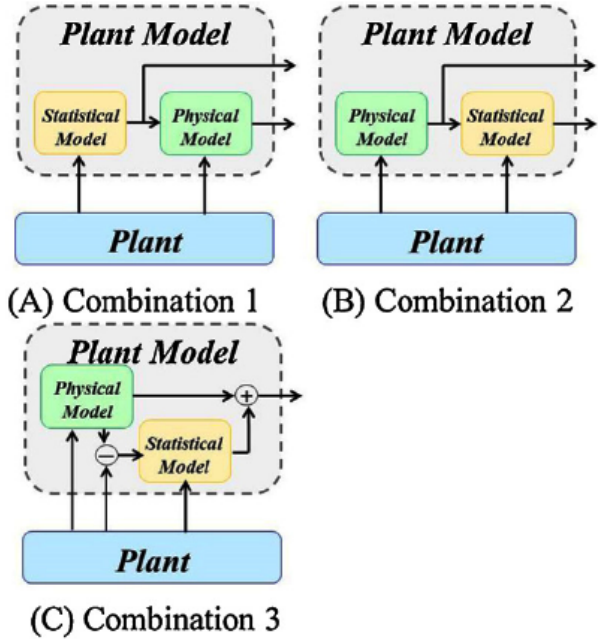

Figure 3: Integrated application of statistical and white box models [32]

The objective function is the same as in the previous case (*Eq. (1)*). The difference is that constraints are nonlinear equations. This problem can be reduced to the linear case. This method is referred to as successive linearisation: nonlinear constraints can be linearised around working points using first-order Taylor series. The solution of the linearisation can be written as

$$\mathbf{h} = \mathbf{B}_y\check{\mathbf{y}} \; + \mathbf{B}_z\check{\mathbf{z}} - \mathbf{f}(\check{\mathbf{y}}_i, \check{\mathbf{z}}_i). \tag{18}$$

The initial values are usually the raw values. After linearisation the procedures of data reconciliation is described in subsection Linear data reconciliation (Q-R factorization, reconciliation of measured values, then estimation of unmeasured variables). If the reconciled and initial values are far from each other, new iteration begins. The initial value will be the result of the previous iteration. Different criteria can be defined: number of iteration or $||\hat{\mathbf{y}}_n - \hat{\mathbf{y}}_{n-1}||$ is smaller than the specified tolerance values.

Successive linearisation is a relatively simple and fast solution, but variable bounds cannot be handled with this method. Another option to solve nonlinear data reconciliation problem is using Nonlinear Programming (NLP) techniques that can estimate measured and unmeasured variables simultaneously. Sequential quadratic programming (SQP) and generalized reduced gradient (GRG) are usual techniques in handling nonlinear problems. These methods are more computationally demanding, but they are numerically more robust [33].

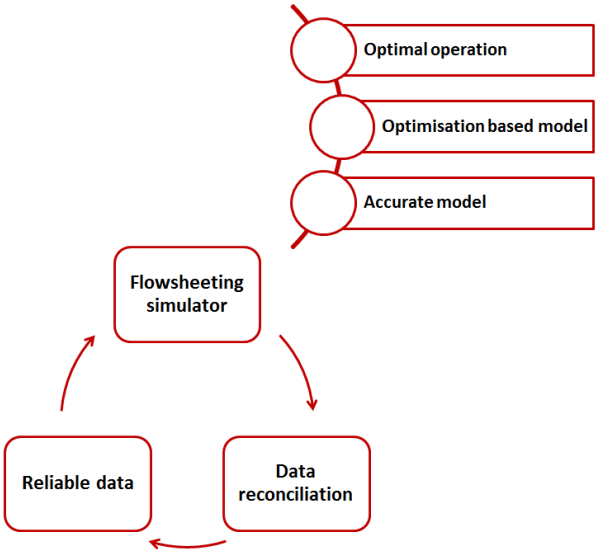

Figure 4: Iterative model development

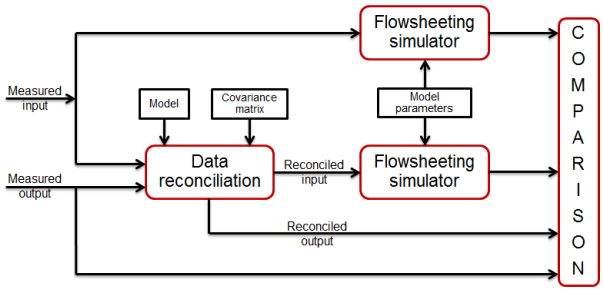

Figure 5: Options of data processing in flowsheeting simulator

Combined application of data reconciliation and flowsheeting simulator

Mathematical models describe the connection between the relevant properties of the analyzed phenomena, process, or activities. In the case of complex chemical processes, mathematical models usually contain differential equations whose solution is not always possible analytically. The process simulator (it is often applied to improve complex industrial processes) contains models of the equipments so their mathematical description is unnecessary for the users. In addition, it includes different thermodynamic models and component databases. The flowsheeting simulator can be defined as "Use of a computer program to quantitatively model characteristic equations of a chemical process" [34]. The simulator is used in batch processing, integrated process engineering (for example, economic analysis and supply chain forecast, supervisory process control, on-line modeling and optimization, safety and reliability analysis), and process synthesis and design (for instance, heat integration, conceptual design). However, lots of information are needed from technology for the simulator to describe the real process exactly. Many of the needed data are measured (e.g., flow, composition, temperature, pressure). If the input of the simulator is faulty, the simulator cannot give reliable results.

A new direction is the combination of simulators and data based statistical models that can handle random and gross errors [23, 35]. The structure of these hybrid models are shown in *Fig. 3*.

- In the first case, the statistical model is the input of the physical model in the form of differential or algebraic equations or a complex flowsheeting simulator. In this case, the statistical model is used to estimate parameters and phenomena that are difficult to model.
- Combination 2 shows the case when outputs of the physical model are transformed by a statistical model.
- In the third option, the difference between measured and calculated variables are the inputs of the statistical model used for correction.

Model-based data reconciliation techniques are similar to the third approach. Using data reconciliation technique random errors can be filtered. When accurate data are available, the flowsheeting simulator gives more reliable results that make further development of models possible. This means that data reconciliation allows us to check the reliability of these measurements [36] and reconsolidated data can be used to build accurate models.

The developed model based data reconciliation technique

Our goal is to design an expert system that can be used to check the acceptability of measurements and improve the flowsheeting simulator of technology. The proposed method is based on model-based data reconciliation techniques. Using data reconciliation, faulty measurements can be found. Random errors can be filtered using discrete Fourier transform. The frequency of the noise is much higher than the frequency of basic process so the discrete Fourier transform gives a graphical view about the differences between the basic process and the noise. Thus, the filtration of the frequency can be determined so the noise can be separated from the basic process. If the amplitude of the noise is known, the standard deviation can be estimated. Thereby, input of simulator is error-free, so systematic model mismatch can be recognized and the tuning of the model can be initiated. Data reconciliation requires an accurate model whose parameters require reconciled process data. In th case of an efficient iterative procedure, a properly working simulator can be developed (see *Fig. 4*).

Since the simulator is able to process historical measured data, the difference between the theoretically achievable values and the measured outputs can give useful information (see *Fig. 5*). Four different output

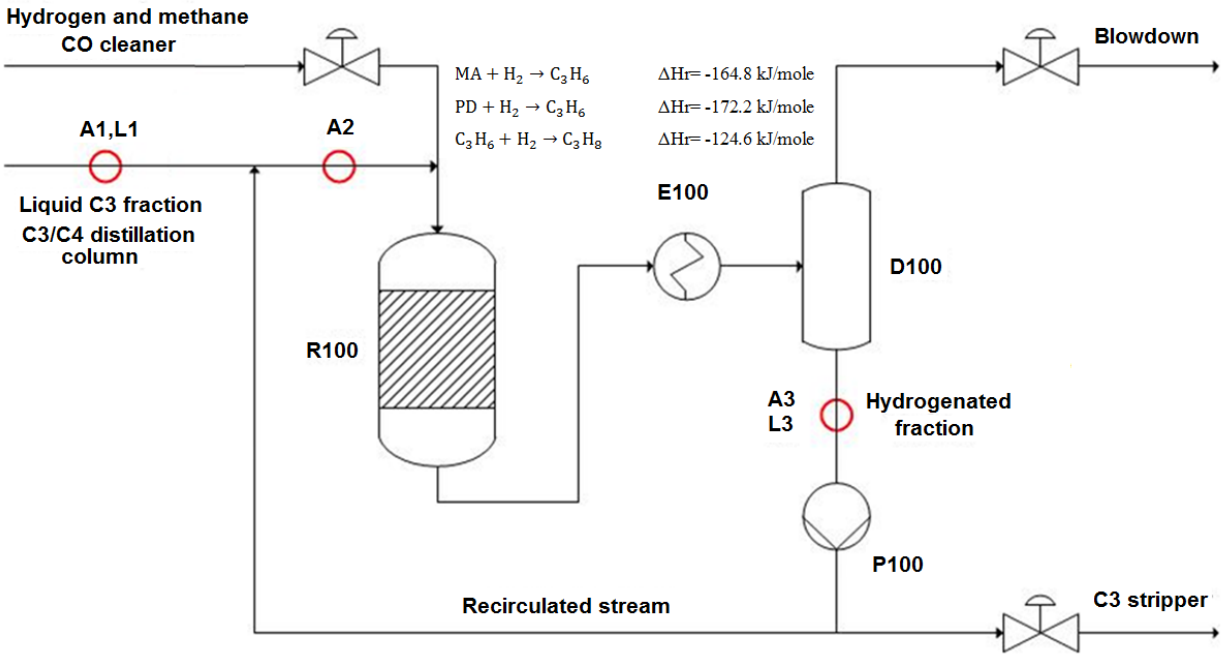

Figure 6: Flowsheet diagram of the C3 hydrogenation system

values can be compared in the case of the same variable. Measured output values are known. If all necessary information is available for data reconciliation, we get the reconciled output value. In addition, if process model is prepared in the flowsheeting simulator, we have two options for analysis: the input of the simulator can be the measured or the reconciled value.

The difference between outputs can be used for two purposes. If the difference is insignificant, the measured data is acceptable and measuring instruments operate properly; calibration is not necessary. Moreover, it helps in the development of the flowsheeting simulator. If the calculated results with reconciled input are far from the reconciled output, the simulator do not describe the real process properly; maybe a parameter is not accurate or it needs structural change.

The proposed method is illustrated on the basis of an industrial hydrogenation system. The case study shows examples for both cases; when the difference indicates the conformity of measurements and when it warns that some model parameters are incorrect.

Case study

The Tisza Chemical Group Plc. (TVK) is the largest petrochemical company of Hungary where polymer raw materials (ethylene, propylene, butylenes, etc.) are produced by steam cracking of naphtha or gasoline. At high temperature, numerous free radical reactions occur. Cracked gas includes light components that are produced in larger amounts, for example, methane, ethane, ethylene, acetylene, propane, propylene, methyl-acetylene, propadiene, n-butane, isobutane, 1,3-butadiene, etc. Some of them are undesirable, because they poison the catalysts for polymerization reactions (e.g. acetylene, methyl-acetylene, propadiene). Another problem is that separation of these components is difficult from the main products by distillation. Due to these problems, undesirable hydrocarbons are hydrogenated. Hydrogenation process of methyl-acetylene and propadiene is presented in this case study (see *Fig. 6*).

The C3-selective hydrogenation process transforms methyl-acetylene and propadiene to propylene by catalytic reaction avoiding the transformation of propylene to propane. The methyl-acetylene and propadiene content of hydrogenated C3 fraction should be less than 1000 ppm. The concentration of the main components (methyl-acetylene, propane) is measured by online analysers (A1, A2 and A3) and there are two places where sample is taken twice a day for laboratory analysis (L1 and L3).

Hydrogen and liquid olefin stream are fed to the reactor where chemical reactions take place. Outlet stream of the reactor is cooled by water so the gaseous C3 component condenses and the hydrogen and the liquid olefin phases are separated. The recirculated stream has two important functions: cools the reactor and dilutes the inlet C3 stream. Blowdown is required due to the accumulation of inert components.

The most frequent type of hydrogenation reactors is the trickle-bed reactor in the olefin plant. Liquid olefin feed and gaseous hydrogen pass through the catalyst bed in the same direction: from top to bottom. Cooling of reactor is provided by vaporization liquid flow. The reactions occur in the top region of the reactor. The

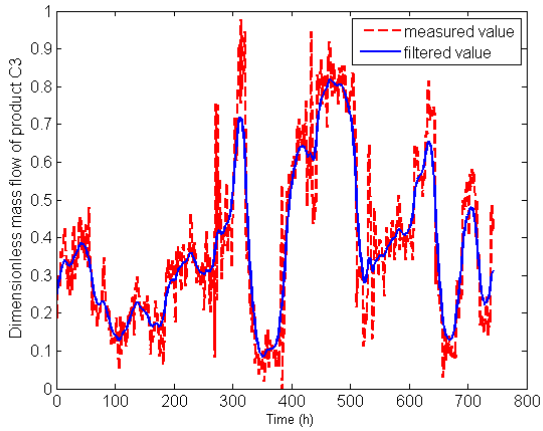

Figure 7: Measured and filtered values of product C3 fraction

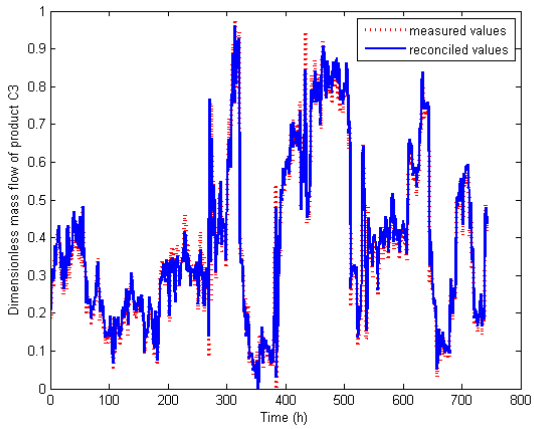

Figure 8: Measured and reconciled values of product C3 fraction

reactions release a large amount of heat. The heat generation is influenced by the methyl-acetylene and propadiene (MAPD) content of the inlet C3 and by the partial pressure of the hydrogen. The MAPD content of the inlet flow of system is high, so the inlet flow of the reactor is diluted. In the range of $10 - 80\,^{\circ}\mathrm{C}$ temperature, the average reaction heat of main reactions are the following:

$$\mathrm{MA} + \mathrm{H_2} = \mathrm{C_3H_6} \quad \Delta H_\mathrm{r} = -164.8\,\mathrm{kJ/mole} \tag{19}$$

$$\mathrm{PD} + \mathrm{H_2} = \mathrm{C_3H_6} \quad \Delta H_\mathrm{r} = -172.2\,\mathrm{kJ/mole} \tag{20}$$

$$\mathrm{C_3H_6} + \mathrm{H_2} = \mathrm{C_3H_8} \quad \Delta H_\mathrm{r} = -124.6\,\mathrm{kJ/mole} \tag{21}$$

In the following sections, released heat vaporizes a part of the liquid phase. This phenomenon also ensures the cooling of the product flow.

The simulator has been developed in Aspen Plus software as part of an expert system used for the monitoring and qualification of the operation of the technology. Calculations provide useful information about unmeasured variables, validity of on-line analyzers, and efficiency of catalysts. The residence time is small (less than one minute). The sample time was five minutes, so the steady state simulation gives the same result as the dynamic simulation. Historical process data entered the simulator by Aspen Simulation Workbook. The developed Excel and Visual Basic macro based framework allows the comparison of the measured data and the calculated results.

Results and discussion

The process has two inlet streams (C3 fraction and hydrogen) and two outlet streams (hydrogenated liquid C3 fraction and blowdown). Mass flows of every stream are measured. Based on the law of conservation of mass (there is no accumulation):

$$\text{Inlet C3} + \mathrm{H_2} = \text{Blowdown} + \text{Product C3} \tag{22}$$

If we check whether the measured data satisfy the balance equation, we find that *Eq. (22)* is not fulfilled, so data reconciliation is necessary.

Steps of the procedure are introduced based on mass flow of C3 product fraction. First time random error is separated from basic process using discrete Fourier transformation. *Fig. 7* shows the measured and the filtered values of product C3 fraction. We defined the random noise as the difference between the measured and filtered values so the standard deviation and covariance matrix of random error ($\mathbf{V}_{\tilde{\mathbf{d}}}$) can be calculated.

Since there is no accumulation (i.e., $b = 0$) *Eq. (10)* is simplified to

$$\hat{\mathbf{y}} = \left(\mathbf{I} - \mathbf{V}_{\tilde{\mathbf{d}}}\mathbf{A}^T\left(\mathbf{A}\mathbf{V}_{\tilde{\mathbf{d}}}\mathbf{A}^T\right)^{-1}\mathbf{A}\right)\tilde{\mathbf{y}}. \tag{23}$$

Since two input and two output streams are in the system, the incidence matrix is $\mathbf{A} = [1\ 1\ \ -1\ \ -1]$. Every needed information are known so the matrix operation can be performed. Real industrial data are analyzed, so dimensionless units are included in the figures. The result of reconciliation is visible in *Fig. 8*. The result shows that the proposed method gives a minor improvement of the process values. Thanks to the validated accuracy of the mass measurements, these process values can be directly used for the validation of the simulator.

In the next step, we compare the reconciled output with the values calculated by the simulator in two different ways. First, the input of the simulator is the measured data, then, the reconciled input data. The result of the comparison is shown in *Fig. 9*. First, we thought that the three curves coincide. However, as a portion is zoomed, the difference became apparent (see *Fig. 10*). If the input of simulator is the reconciled data, the calculated and the reconciled output are nearly the same. Although the difference between the two calculated curves may seem small in dimensionless space, do not forget that the small percentage difference can mean significant mass flow values in a real industrial process.

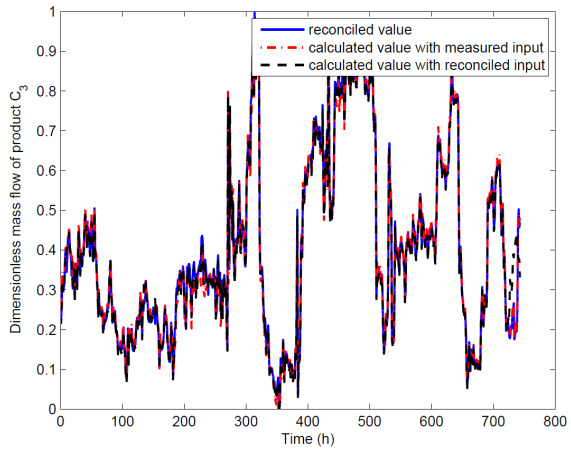

Figure 9: Calculated and reconciled data of product C3 fraction if the input of the simulator is the measured and reconciled data

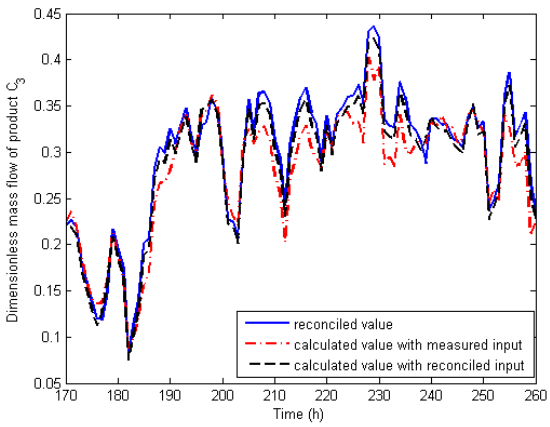

Figure 10: Reconciled and calculated outputs of simulator in case of product C3 fraction

We calculated the square error in both cases (deviation from the reconciled output) and found that the difference decreased by 87% if the reconciled values are the input of the simulator. This shows that the simulator gives more accurate results if previously data are reconciled.

The simulator supported by data reconciliation can be applied in fault diagnosis. *Fig. 11* shows how the hydrogen flow changes over time. The continuous line represents the simulator results without data reconciliation, while the dashed line shows the reconciled values.

Fig. 11 shows the similarity of the calculated and the measured values. After 300 hours of operation the calculated values exceed the measured data. There are several possible reasons. Less hydrogen goes to this system if

- The MAPD content of the inlet stream is lower, because the mass flow of hydrogen is controlled in proportion of the MAPD mole flow of the input C3 fraction.
- After the regeneration when catalyst is active.
- Other reactor is regenerated or reactivated.

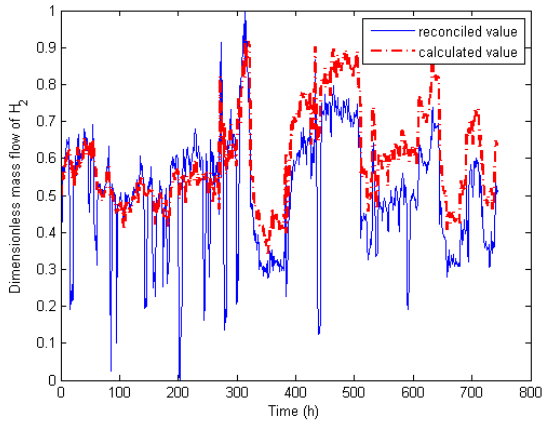

Figure 11: Calculated and reconciled dimensionless mass flow of H_2 if the input of the simulator is the measured data

In this case, another reactor in the Olefin 2 plant was reactivated (after 300 hours). The reactivation process needed a lot of hydrogen so the hydrogen for the studied reactor was reduced. The received data did not include the rate of reduction. The result of the simulation showed that something was changed in the technology. This shows the technique can be used for monitoring and diagnostics of complex processes because it can be discovered if the operating condition has changed.

Conclusion

Data reconciliation is an effective method to obtain accurate process data that satisfy balance equations using process constraints. The application of steady-state data reconciliation is widespread in the process industry. Simulators are often applied to improve industrial processes, optimize operation, and identify bottlenecks of technology. Historical process data can be used for the identification and verification of models utilized by these tools. Usually, measured data do not satisfy balance equations, because all measurements are incorrect to some extent. Thus, it is necessary to develop a method which can simultaneously and iteratively improve data and model performance. A method based on data reconciliation technique has been developed for this purpose.

An industrial C3 selective hydrogenation process has been analysed as a case study. The steady state simulator of the process has been developed in Aspen Plus flowsheeting software. Calculated and historical process data were compared using Aspen Simulation Workbook.

The proposed technique can be used for monitoring complex processes. Results showed that the proposed technique slightly improves the quality of flow data. Thanks to the accuracy of flow measurements, these process values are directly applicable for the validation of simulator and suitable for further studies, e.g., for the determination of kinetic parameters.

Acknowledgements

This research of Barbara Farsang and Janos Abonyi was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program'. The infrastructure of the research has been supported by the European Union and the Hungarian Republic through the TÁMOP 4.2.2/A-11/1/KONV-2012-0071 project.

SYMBOLS

$\mathbf{A}$	incidence matrix
$\mathbf{A}_y$	incidence matrix of balances in terms of measured flows
$\mathbf{A}_z$	incidence matrix of balances in terms of unmeasured flows
$\mathbf{B}_y$	Jacobian matrix of measured variables
$\mathbf{B}_z$	Jacobian matrix of unmeasured variables
$\mathbf{b}$	source of extensive quantity
$\mathbf{d}$	$\mathbf{y} - \check{\mathbf{y}}$ error
$\mathbf{f}$	function vector of equality model constraints
$\tilde{\mathbf{f}}$	balance error ($\mathbf{A}\tilde{\mathbf{y}} - \mathbf{b}$ vector)
$\hat{\mathbf{f}}$	balance error ($\mathbf{A}\hat{\tilde{y}} - \mathbf{b}$ vector)
$\mathbf{g}$	function vector of inequality model constraints
$\mathbf{I}$	identity matrix
MA	methyl-acetylene
MAPD	methyl-acetylene and propadiene
PD	propadiene
$\mathbf{P}$	projection matrix
$\mathbf{V}_{\tilde{\mathbf{d}}}$	covariance matrix of measurements
$\check{\mathbf{y}}$	vectors of real values of measured variables
$\tilde{\mathbf{y}}$	vectors of measured variables
$\hat{\mathbf{y}}$	vectors of reconciled measured variables
$\check{\mathbf{z}}$	vectors of real values of unmeasured variables
$\tilde{\mathbf{z}}$	vectors of unmeasured variables
$\hat{\mathbf{z}}$	vectors of estimated unmeasured variables
λ	Lagrange multiplier
ΔH_r	reaction heat

REFERENCES

[1] BAUER, M., CRAIG, I. K. Economic assessment of advanced process control - A survey and framework, Journal of Process Control, *2008*, 18(1), 2–18

[2] PACH, F. P. , FEIL, B. , NÉMETH, S., ÁRVA, P. , ABONYI, J. Process data warehousing based operator support system for complex production technologies, IEEE Transactions on Systems, Man and Cybernetics, Part A, Special Issue on 'Advances in Heterogeneous and Complex System Integration', *2006*, 136–153

[3] TIETZE, U., SCHENK, C. Analóg és digitális áramkörök (Műszaki Könyvkiadó)

[4] KUHEN, D., DAVIDSON, H. Computer Control II. Mathematics of Control, Chemical Engineering Progress, *1961*, 57, 44–47

[5] NARASIMHAN, S. Introduction to data reconciliation and gross error diagnosis (2013), URL: www.che.iitm.ac.in/ naras/ ch544/introDRGED.pdf

[6] BELLEC, S., JIANG, T., KERR, B., DIAMOND, M., STUART, P. On-line processing and steady state reconciliation of pulp and paper mill process data, 91st Annual Meeting Preprints-Book B, *2005*, 105–110

[7] THIBAULT, J. Introduction to Data Reconciliation *2013*, URL: www.polymtl.ca/namp/docweb/ Modules_Web/M11_Tier1_Chap1-3.pdf

[8] NOGITA, S. Statistical Test and Adjustment of Process Data, Ind. Eng. Chem, Process Des. Dev., *1972*, 11, 197–200

[9] MAH, R. S. H., STANLEY, G. M., DOWNING, D. M. Reconciliation and Rectification of Process Flow and Inventory Data, Ind. Eng. Chem, Process Des. Dev., *1976*, 15, 175–183

[10] MAH, R. S. H. Chemical process structures and information flows, AIChE Journal, *1993*, 39(4), 730

[11] CAO, S., RHINEHART, R. An efficient method for on-line identification of steady state, Journal of Process Control, *1995*, 5(6), 363–374

[12] TJOA, I. B., BIEGLER, L. T. Simultaneous solution and optimization strategies for parameter estimation of differential – algebraic equation systems, Industrial and Engineering Chemistry Research, *1991*, 30, 376–385

[13] BRITT, H. I. , LUECKE, R. H. The estimation of parameters in nonlinear implicit models, Technometrics, *1973*, 15, 233–247

[14] MADRON, F., VEVERKA, V. Optimal selection of measuring points in complex plants by linear models, American Institute of Chemical Engineering Journal, *1992*, 38(2), 227

[15] CROWE, C. M. Reconciliation of process flowrates by matrix projection. Part II: The nonlinear case, AIChE Journal, *1986*, 32, 616–623

[16] BAGAJEWICZ, M. J., JIANG, Q. Integral approach to plant linear dynamic reconciliation, AIChE Journal, *1997*, 43, 2546–2558

[17] PRATA, D. M., LIMA, E. L., PINTO, J. C. Simultaneous data reconciliation and parameter estimation in bulk polypropylene polymerizations in real time, Macromolecular Symposia, *2006*, 243, 91–103

[18] PRATA, D. M., SCHWAAB, M., LIMA, E. L., PINTO, J. C. Nonlinear dynamic data reconciliation and parameter estimation through particle swarm optimization: Application for an industrial polypropylene reactor, Chemical Engineering Science, *2009*, 64 18, 3953–3967

[19] MARTÍNEZ-MARADIAGA, M., BRUNO, J. C., CORONAS, A. Steady-state data reconciliation for absorption refrigeration systems, Applied Thermal Engineering, *2013*, 51(1-2), 1170–1180

[20] SARABIA, D., DE PRADA, C., GOMEZ, E., GUTIERREZ, G., CRISTEA, S.,SOLA, J., GONZALEZ, R. Data reconciliation and optimal management of hydrogen networks in a petrol refinery, Control Engineering Practice, *2012*, 20(4), 343–354

[21] CHEN, J., PENG, Y., MUNOZ, J. C. Correntropy estimator for data reconciliation, Chemical Engineering Science, *2013*, 104, 1019–1027

[22] KELLER, T., HOLTBRUEGGE, J., GÓRAK, A. Transesterification of dimethyl carbonate with ethanol in a pilot-scale reactive distillation column, Chemical Engineering Journal, *2012*, 180, 309–322

[23] DOMINGUES, L:, PINHEIRO, C. C. , OLIVEIRA, N. M., VILELAS, A., FERNANDES, J., RIBEIRO, F. R. Estimation of catalyst deactivation parameters of ethyl tert-butyl ether (ETBE) reactors based on industrial plant data, in 22nd European Symposium on Computer Aided Process Engineering, *Computer Aided Chemical Engineering*, vol. 30 (I. D. L. BOGLE, M. FAIRWEATHER, eds.) (Elsevier) *2012*, pp. 1002–1006

[24] PRATA, D. M. , SCHWAAB, M., LIMA, E. L., PINTO, J. C. Simultaneous robust data reconciliation and gross error detection through particle swarm optimization for an industrial polypropylene reactor, Chemical Engineering Science, *2010*, 65(17), 4943–4954

[25] VASEBI, A., POULIN, E., HODOUIN, D. Dynamic data reconciliation in mineral and metallurgical plants, Annual Reviews in Control, *2012*, 36(2), 235–243

[26] VALDETARO, E. D., SCHIRRU, R. Simultaneous Model Selection, Robust Data Reconciliation and Outlier Detection with Swarm Intelligence in a Thermal Reactor Power calculation, Annals of Nuclear Energy, *2011*, 38(9), 1820–1832

[27] MARTINS, M. A., AMARO, C. A., SOUZA, L. S., KALID, R. A. , KIPERSTOK, A. New objective function for data reconciliation in water balance from industrial processes, Journal of Cleaner Production, *2010*, 18(12), 1184–1189

[28] STANLEY, G., MAH, R. Observability and redundancy in process data estimation, Chemical Engineering Science, *1981*, 36(2), 259–272

[29] CROWE, C. M. Observability and redundancy of process data for steady state reconciliation, Chemical Engineering Science, *1989*, 44(12), 2909–2917

[30] ALI, Y., NARASIMHAN, S. Redundant sensor network design for linear processes, American Institute of Chemical Engineering Journal, *1995*, 41(10), 2237

[31] ALMASY, G.: Mérlegegyenletek - mérési hibák (Egyetemi Kiadó, Veszprém) *1996*

[32] NAKAYA, M., LI, X. On-line tracking simulator with a hybrid of physical and Just-In-Time models, Journal of Process Control, *2013*, 23(2), 171–178

[33] THIBAULT, J. tate Data reconciliation for Bilinear Systems, *2013*, URL: www.polymtl.ca/namp/docweb/ Modules_Web/M11_Tier1_Chap4-5.pdf

[34] FLORES, A.: Introduction to Aspen Plus (12. 09. 2001)

[35] KANOA, M., NAKAGAWA, Y. Data-based process monitoring, process control and quality improvement: Recent developments and applications in steel industry, Computers and Chemical Engineering, *2008*, 32, 12–24

[36] KADLEC, P., GABRYS, B., STRANDT, S. Data-driven Soft Sensors in the process industry, Computers and Chemical Engineering, *2009*, 33, 795–814