ELSEVIER

ISA Transactions 39 (2000) 35–45

ISA TRANSACTIONS®

www.elsevier.com/locate/isatrans

Data reconciliation with uncertain models

Didier Maquin*, Olivier Adrot, José Ragot

Centre de Recherche en Automatique de Nancy, CNRS UPRES-A 7039, 2, Avenue de la forêt de Haye, 54516 Vandoeuvre Cedex, France

Abstract

Data reconciliation has proven to be an effective technique for providing frequent, accurate and consistent "best estimates" of plant operation data. However, in almost all the proposed techniques until today, the mathematical model of the process has been considered as exact. In point of fact, this hypothesis is uncommon and frequently the models used are uncertain. This paper proposes a new technique of data reconciliation which is able to exploit the knowledge about the uncertainties of the model with regard to which the reconciliation is done. It leads to the solution of a classical quadratic optimisation problem subject to constraints. The originality of the proposed technique is to use penalty functions for solving this problem and to weight each constraint with regard to their uncertainties. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Data reconciliation; Uncertain models; Estimation techniques

1. Introduction

Data reconciliation is a technique to improve the quality of measured plant operation data. These measurements are obtained using instruments that are inherently inaccurate or subject to failures. Using erroneous data for technical analysis and decision-making may yield distorted conclusions and result in improper decisions. Accurate data are, therefore, essential for monitoring, analysing and controlling plant operation. One measure of data inaccuracy is the consistency of the data with regard to the mathematical models describing the considered process. Among the more classical models used for describing the functioning of a process are the balance relationships (mass, component, species, enthalpy, chemical thermodynamic, etc.). If all of these models are structurally perfectly known, some of them depend on parameters which are difficult to assess. Therefore, it becomes very hazardous and mathematically not correct to reconcile operation data with regard to an uncertain model without taking this last fact into account. Assuming some knowledge about the precision of the values of the parameters, we propose to take this information into account in the reconciliation procedure. Some attempts in this direction have already been recently presented [1,2]. The proposed technique is limited, in this paper, to the case of linear models (with regard to variables and with regard to parameters).

Let us consider a process characterised by the vector of the true variables X^* (of dimension ν) and represented by the linear model:

$$M(\theta)X^* = 0 \tag{1a}$$

* Corresponding author. Tel.: +33-3-83-59-59-59; fax: +33-3-83-59-56-44.

E-mail address: dmaquin@ensem.u-nancy.fr (D. Maquin).

0019-0578/00/$ - see front matter © 2000 Elsevier Science Ltd. All rights reserved.
PII: S0019-0578(99)00045-2

where $M(\theta)$, of dimension $n \times v$, is a matrix depending on a parameter vector θ for which the mathematical expectation θ_0 and the diagonal variance matrix W are assumed to be known. This formulation notably allows the process to be described with material balance equations. The representation (1a) is, however, general and can as well be applied to dynamic discrete systems or discretised continuous dynamic systems [3]. The measures X of the variables X^* are corrupted by errors that are supposed to be distributed as Gaussian random variables, with zero mean and a known variance-covariance matrix V; the measures are linked, in an additive way, to the real values X^* and to the measurement errors ε by the relation:

$$X = X^* + \varepsilon \tag{1b}$$

Formulation (1b) states that all variables are measured. If it is not the case, a classical observability analysis may extract the redundant part on which the proposed treatment can be applied [4]. An estimation of the state $\hat{X}$ of the process may be obtained by minimising, under the constraints (2b), the quadratic function (2a):

$$\phi = \frac{1}{2}\left\|\hat{X} - X\right\|^2_{V^{-1}} \tag{2a}$$

$$M(\theta)\hat{X} = 0 \tag{2b}$$

The choice of a quadratic function may be easily justified when the errors are normally distributed [5]. The aim of this paper is to propose a simple method for solving problem (2). The whole paper is organised as follows. After this introduction, describing the problem and enlightening the difficulty for accounting model uncertainties, the second section is devoted to the recalling of the solution to the problem of data reconciliation for linear systems with regard to a perfectly known model. The obtained solution is analytical. The principles of utilisation of penalty functions for solving a quadratic optimisation problem subject to linear constraints are presented in the third section. Although sub-optimal, this method provides degrees of freedom for taking into account the model uncertainties as it will be shown in Section 5. Then, it is shown that the solution of the previous problem may be formulated in a recurrent fashion. Based on the use of penalty functions, a methodology to calculate the weighting factor of each model equation in order to take into account their uncertainty degrees is suggested in the sixth section. The seventh section is dedicated to the problem of sensor fault detection. It is shown how the proposed technique enhances the pertinence of decision-making stage. At last, the method is applied on a small didactic example.

2. Optimal solution

Let us first consider the problem (2) with $\theta =$ constant. In this case, we denote $M(\theta)$ as M. Then, the problem to be solved is a quadratic optimisation problem subject to linear constraints. Its optimal solution may be obtained by applying the technique of Lagrange multipliers. The associated Lagrangian may be written as:

$$L = \phi + \lambda^T M\hat{X} \tag{3}$$

where λ is the n-dimensional vector of Lagrange parameters. The solution is given by the first order stationary conditions of this Lagrangian. It yields the following $n + v$ simultaneous equations:

$$\frac{\partial L}{\partial \hat{X}} = V^{-1}\left(\hat{X} - X\right) + M^T\lambda = 0 \tag{4a}$$

$$\frac{\partial L}{\partial \lambda} = M\hat{X} = 0 \tag{4b}$$

From this system, the following solution may be easily deduced:

$$\hat{X} = P_r X \tag{5a}$$

where P_r is a so-called projection matrix defined by:

$$P_r = I - VM^T(MVM^T)^{-1}M \tag{5b}$$

Notice that the inverse of the matrix MVM^T always exists provided that all the constraints (rows of M) are independent. Knowing that the vector of measurement errors ε is normally distributed with

a zero mean and a known diagonal variance matrix V, let us now demonstrate that the obtained estimator is unbiased, indeed, we have:

$$\mathrm{E}(\hat{X}) = \mathrm{E}(P_r(X^* + \varepsilon)) = P_r\mathrm{E}(X^*) + P_r\mathrm{E}(\varepsilon) \tag{6}$$

as $\mathrm{E}(X^*) = X^*$ and $\mathrm{E}(\varepsilon) = 0$, we obtain:

$$\mathrm{E}(\hat{X}) = P_r X^* = X^* \tag{7}$$

because, of course, the true values satisfy the constraints. The variance-covariance matrix of the estimator may now be evaluated.

$$\mathrm{Var}(\hat{X}) = \mathrm{E}\Big((\hat{X} - \mathrm{E}(\hat{X}))(\hat{X} - \mathrm{E}(\hat{X}))^T\Big) \tag{8}$$

From Eqs. (6) and (7), it yields:

$$\hat{X} - \mathrm{E}(\hat{X}) = P_r(X^* + \varepsilon) - X^* = P_r\varepsilon \tag{9}$$

Then, Eq. (8) may be written as:

$$\mathrm{Var}(\hat{X}) = \mathrm{E}\big(P_r\varepsilon(P_r\varepsilon)^T\big) = P_r\mathrm{E}\big(\varepsilon\varepsilon^T\big)P_r^T = P_r V P_r^T \tag{10}$$

And after some simple matrix calculus, one obtains:

$$\mathrm{Var}(\hat{X}) = \hat{V}_x = P_r V \tag{11}$$

Summarising, Eqs. (5) and (11) characterise the optimal solution of the problem. It may be noticed that the vector of estimates and the corresponding variance-covariance matrix are linked to the vector of measurements and its variance matrix by the same projection matrix P_r. The enhancement of the quality of the measurements may be evaluated by comparing the diagonal terms of V and $\hat{V}_x$.

3. Sub-optimal solution

A very classical solution for solving optimisation problem subject to constraints is to use penalty functions. This method leads to a sub-optimal solution which, however, can be as close as desired to the optimal one. Let us consider again the problem (2). The previous optimisation criterion is modified as follows:

$$\phi_m = \frac{1}{2}\left(\left\|X - \hat{X}\right\|^2_{V^{-1}} + k^2\left\|M\hat{X}\right\|^2\right) \tag{12}$$

where k is a scalar value which weights the constraints. The problem to be solved is then reduced to a quadratic optimisation problem without constraints. Due to the additive form of Eq. (12), it is clear that the optimal state $\hat{X}$ is the result of a compromise between two objectives; the first one being the respect of the statistical distribution of the measurements errors and the second corresponding to the constraints satisfaction. Intuitively, one can see that the constraints will only be satisfied if the value of k tends to infinity. This remark will be demonstrated further. The solution which minimises the value of the criterion (12) is obtained from the following equation:

$$\frac{\partial\phi_m}{\partial\hat{X}} = V^{-1}\Big(\hat{X} - X\Big) + k^2M^TM\hat{X} = 0 \tag{13}$$

Therefore, one obtains:

$$\hat{X} = \big(V^{-1} + k^2M^TM\big)^{-1}V^{-1}X \tag{14}$$

Using the lemma of matrix inversion, this previous expression may also be written as follows:

$$\hat{X} = PX \tag{15a}$$

with:

$$P = I - VM^T\big(k^{-2}I + MVM^T\big)^{-1}M \tag{15b}$$

So, when the scalar value k tends to infinity, the obtained solution is identical as that given by (5). For all the other values of k, the estimation given by (15) does not satisfy the constraints. The corresponding a posteriori residual, which is a decreasing function of the weight k is given by:

$$\hat{R} = M\hat{X} = M\big(V^{-1} + k^2M^TM\big)V^{-1}X \tag{16}$$

As before, the statistical properties of this estimator may be calculated. Following the same method, we already have:

$$\mathrm{E}(\hat{X}) = PX^* \tag{17}$$

that is to say:

$$\mathrm{E}(\hat{X}) = \left(I - VM^T\left(k^{-2}I + MVM^T\right)^{-1}M\right)X^* \quad (18)$$

Using the property $MX^* = 0$, it leads to:

$$\mathrm{E}(\hat{X}) = X^* \quad (19)$$

Therefore, this new estimator is also unbiased and its variance–covariance matrix is equal to:

$$\mathrm{Var}(\hat{X}) = PVP^T = (I + k^2VM^TM)^{-2}V \quad (20)$$

In this case, the reduction of the variance of the estimate with regard to the variance of the measurements depends on the chosen weight k.

4. Recurrent formulation of the sub-optimal solution

The proposed solution may be extended to the case of multiple penalty functions. Let us denote k_j^2, the weighting factor of the jth constraint in the optimisation criterion. This latter becomes:

$$\phi_m = \frac{1}{2}\left(\left\|X - \hat{X}\right\|_{V^{-1}}^2 + \sum_{j=1}^{n} k_j^2 \left\|m_j\hat{X}\right\|^2\right) \quad (21)$$

It corresponds to the situation where it is desirable to take into account, separately, the "quality" of each constraint. This point will be analysed in Section 5. The solution of this problem may be obtained recurrently by adding, successively, each constraint. In order to explain this procedure, let us first consider the following partition of the constraint matrix:

$$M = \begin{pmatrix} M_{n-1} \\ m_n \end{pmatrix} \quad (22)$$

where m_n represents the last constraint (m_n is a ν-dimensional vector). The optimisation criterion may be written as follows:

$$\phi_m = \frac{1}{2}\left(\left\|X - \hat{X}\right\|_{V^{-1}}^2 + \left\|M_{n-1}\hat{X}\right\|_{K_{n-1}^2}^2 + k_n^2\left\|m_n\hat{X}\right\|^2\right) \quad (23)$$

The matrix K_{n-1}^2 is diagonal and contains the weights associated with the different constraints of M_{n-1}.

The minimum of the criterion (23), with regard to $\hat{X}$, is obtained when:

$$\frac{\partial\phi_m}{\partial\hat{X}} = V^{-1}(\hat{X} - X) + M_{n-1}^T K_{n-1}^2 M_{n-1}\hat{X} + k_n^2 m_n^T m_n\hat{X} = 0 \quad (24)$$

whence:

$$\hat{X} = \left(V^{-1} + M_{n-1}^T K_{n-1}^2 M_{n-1} + k_n^2 m_n^T m_n\right)^{-1} V^{-1}X \quad (25)$$

Using the lemma of matrix inversion, this previous expression may also be written as follows:

$$\hat{X} = \left(Y_{n-1} - Y_{n-1}m_n^T\left(k_n^{-2} + m_n Y_{n-1} m_n^T\right)^{-1} m_n Y_{n-1}\right) V^{-1}X \quad (26a)$$

with:

$$Y_{n-1} = \left(V^{-1} + M_{n-1}^T K_{n-1}^2 M_{n-1}\right)^{-1} \quad (26b)$$

Therefore, the "global" estimation $\hat{X}$ (26a), obtained with the n constraints, may be written using the "partial" one $\hat{X}_{n-1}$ (those obtained by taking into account the constraints M_{n-1} only):

$$\hat{X} = \left(I - Y_{n-1}m_n^T\left(k_n^{-2} + m_n Y_{n-1} m_n^T\right)^{-1} m_n\right)\hat{X}_{n-1} \quad (27)$$

This procedure may be extended to the particular case where the constraints are progressively taken into account one after the other. The constraint matrix is, therefore, split into n row vectors:

$$M = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{pmatrix}$$

It is, therefore, easy to show that expression (27) may be generalised in order to write the following recurrences:

$$\hat{X}_{i+1} = \left(I - Y_i m_{i+1}^T\left(k_{i+1}^{-2} + m_{i+1} Y_i m_{i+1}^T\right)^{-1} m_{i+1}\right)\hat{X}_i \quad i = 0, \cdots, n-1 \quad (28a)$$

$$Y_{i+1} = \left(I - Y_i m_{i+1}^T \left(k_{i+1}^{-2} + m_{i+1} Y_i m_{i+1}^T\right)^{-1} m_{i+1}\right) Y_i \quad i = 0, \cdots, n-2 \tag{28b}$$

with:

$$\hat{X}_0 = X \text{ and } Y_0 = V \tag{28c}$$

The estimation obtained at iteration $n-1$, when all the constraints have been taken into account, corresponds to the researched solution, $\hat{X} = \hat{X}_n$. Using expression (18), it is also possible to calculate recurrently the variance-covariance matrix of the estimator:

$$\hat{V}_{i+1} = P_{i+1} \hat{V}_i P_{i+1}^T \tag{29a}$$

with:

$$P_{i+1} = I - Y_i m_{i+1}^T \left(k_{i+1}^{-2} + m_{i+1} Y_i m_{i+1}^T\right)^{-1} m_{i+1} \tag{29b}$$

and:

$$\hat{V}_0 = V \tag{29c}$$

The proposed technique is very interesting because, in the matrices P_j which intervene in the solutions [Eqs. (28a), (28b) and (29b)], the matrix which must be inverted is a scalar. Moreover, the progressive adding of constraints allows the contribution to the final solution of each constraint to be quantified.

5. Choice of the weighting factors

As previously said, the parameters intervening in the description of the model of a system are not always perfectly known. Assuming the knowledge of a nominal value of each parameter as well as its standard deviation, it is possible to enhance the quality of the estimation results. Indeed, if the model equations are uncertain, the complete satisfaction of the constraint $M(\theta)\hat{X} = 0$ is not justified. So, let us define a vector of a posteriori residuals $\hat{R}$ whose components are:

$$\hat{r}_i = m_i(\theta)\hat{X} \quad i = 1, \cdots, n \tag{30}$$

where the estimation $\hat{X}$ has been computed sequentially from (28) taking into account the weights k_i associated to each constraint. Thus, it is clear that the a posteriori residuals (30) are related to the weights.

Reciprocally, the magnitude of each residual $\hat{r}_i$ may be indexed on the "quality" of each corresponding constraint. In order to define more precisely this idea, let us introduce the following reduction factor β_i:

$$\hat{r}_i = \beta_i r_i = \beta_i m_i(\theta) X \tag{31}$$

If the knowledge on the values of the parameters intervening in a constraint is very precise, this constraint may be considered as "exact" and has to be satisfied; it is natural to require a very small *a posteriori* residual. It corresponds to the choice of a reduction factor β_i close to zero or even null. Conversely, if this knowledge is less precise, the magnitude of this residual could be more important. However, remember that the proposed procedure tries to enhance the quality of the measurements. So, even if the knowledge on the parameters is very poor, one can choose to keep, for the value of a posteriori residual, that of a priori residual, that is to say $\beta_i = 1$. Notice that this choice is arbitrary and other solutions may be proposed. A function which "distributes" the reduction factors between these two limit values (0 and 1) might be chosen. We propose to calculate this function following the next procedure.

First, let us define the jacobian matrices, with regard to the variables and with regard to the parameters, of the constraint $M(\theta)X^* = 0$:

$$\frac{\partial(M(\theta)X^*)}{\partial(X^*)^T} = M(\theta) \text{ and } G(X^*) = \frac{\partial(\mathrm{M}(\theta)X^*)}{\partial\theta^T} \tag{32}$$

Let us also define V_{rx} the variance–covariance matrix of the a priori residual $R = M(\theta)X$ under the hypothesis that the parameters θ are perfectly known $\theta = \theta_0 = \text{constant}$:

$$V_{rx} = M(\theta_0) V M^T(\theta_0) \tag{33}$$

When considering that both the variables and the parameters are subjected to random errors, this

variance–covariance matrix becomes, assuming an approximation at the first order:

$$V_r = M(\theta_0)VM^T(\theta_0) + G(X^*)WG^T(X^*) \tag{34}$$

where W is the known diagonal variance matrix of the parameters. As the true value X^* is not known, its value is replaced, in Eq. (34), by the measurements X. For a particular residual $r_i = m_i(\theta)X$, the corresponding standard deviations will be denoted:

$$\sigma_{r_i X} = \left(m_i(\theta_0)Vm_i^T(\theta_0)\right)^{1/2} \tag{35}$$

and:

$$\sigma_{r_i} = \left(m_i(\theta_0)Vm_i^T(\theta_0) + g_i(X)Wg_i^T(X)\right)^{1/2} \tag{36}$$

The reduction factor β_i associated to the ith constraint may now be defined, taking into account these standard deviations. We propose the following formula:

$$\beta_i = 1 - \frac{\left(m_i(\theta_0)Vm_i^T(\theta_0)\right)^{1/2}}{\left(m_i(\theta_0)Vm_i^T(\theta_0) + g_i(X)Wg_i^T(X)\right)^{1/2}} \tag{37}$$

When the variances of the parameters tend to infinity (poor knowledge), the reduction factor tends to one and when these variances tend to zero (total knowledge), it tends to zero. In the other cases, the proposed ratio gives an image of the "quality" of the sub-model $m_i(\theta)X^* = 0$.

On the basis of the previous remarks, the question which now may be put is the following : the reduction factors of each constraint having been chosen, which is the set of weighting factors which intervene in the optimisation problem, that will allow the obtaining of a posteriori residuals given by (31).

To answer this question, let us consider again the partition (22) of the constraints matrix. Using Eq. (27), the a posteriori residual associated to the constraint m_n can be expressed as:

$$m_n\hat{X} = \left(1 - m_n Y_{n-1} m_n^T\left(k_n^{-2} + m_n Y_{n-1} m_n^T\right)^{-1}\right) m_n\hat{X}_{n-1} \tag{38a}$$

Using the lemma of matrix inversion, this residual may also be written as:

$$m_n\hat{X} = \left(1 + k_n^2 m_n Y_{n-1} m_n^T\right)^{-1} m_n\hat{X}_{n-1} \tag{38b}$$

with:

$$\hat{X}_{n-1} = \left(I - VM_{n-1}^T\left(K_{n-1}^{-2} + M_{n-1}VM_{n-1}^T\right)^{-1}M_{n-1}\right)X \tag{38c}$$

and:

$$Y_{n-1} = \left(I - VM_{n-1}^T\left(K_{n-1}^{-2} + M_{n-1}VM_{n-1}^T\right)^{-1}M_{n-1}\right)V \tag{38d}$$

This a posteriori residual may also be written as:

$$m_n\hat{X} = \frac{m_n\hat{X}_{n-1}}{1 + k_n^2 A} \tag{39a}$$

with:

$$A = m_n Y_{n-1} m_n^T \tag{39b}$$

Substituting in (39a) the expression (38c) of $\hat{X}_{n-1}$, the residual becomes:

$$m_n\hat{X} = \frac{r_n - D}{1 + k_n^2 A} \tag{40a}$$

where r_n is the a priori residual of the constraint m_n:

$$r_n = m_n X \tag{40b}$$

and:

$$D = m_n VM_{n-1}^T\left(K_{n-1}^{-2} + M_{n-1}VM_{n-1}^T\right)^{-1}M_{n-1}X \tag{40c}$$

So, from Eq. (40a), the weighting factor k_n^2 may be written as:

$$k_n^2 = \frac{r_n - m_n\hat{X} - D}{m_n\hat{X}A} \tag{41a}$$

or, depending on the reduction factor proposed in (31):

$$k_n^2 = \frac{(1 - \beta_n)r_n - D}{\beta_n r_n A} \tag{41b}$$

Notice that the expression of the weighting factor k_n^2 depends on the values of all the other weighting factors k_i^2 $i = 1, \cdots, n-1$ [see expressions (39b) and (40c) of A and D]. On the basis of this remark, a complete algorithm for determining all the weighting factors may be proposed. It consists of two series of iterations. In the first series, each weighting factor k_i^2 is calculated on the basis of the knowledge of the others k_j^2, $j = 1, \cdots, n$ $j \neq i$. During the second series, the first series is repeated until the convergence of the weighting factors is obtained. In fact, the algorithm is stopped when the relative variation of two consecutive estimations of each weighting factor is less than a fixed threshold. So, the algorithm consists of the three following steps.

Step 1: initialisation

$$K = \begin{pmatrix} k_0 & & 0 \\ & \ddots & \\ 0 & & k_0 \end{pmatrix}$$

Without a priori information, all the initial weighting factors may be arbitrarily chosen. The simplest is to assign a unique constant k_0 to all the weighting factors. Whether we want, at this initial step, to privilege the proximity between the measurements and their estimates or the satisfaction of the constraints, this constant k_0 will be chosen respectively close to zero or tending towards infinity.

Step 2: calculus of the weighting factors

For $i = 1$ to n

Step 2a: partitioning of matrices

$$M = \begin{pmatrix} M_{(-i)} \\ M_i \end{pmatrix} \quad K = \begin{pmatrix} K_{(-1)} & 0 \\ 0 & k_i \end{pmatrix}$$

where $M_{(-i)}$ is equal to the matrix of constraints M from which the ith row m_i has been removed and $K_{(-i)}$ is the matrix of the weights from which the ith row and column have been removed.

Step 2b: calculus of the ith weighting factor

Auxiliary matrices:

$$r_i = m_i X$$

$$D_i = m_i V M_{(-i)}^T \left(K_{(-i)}^{-2} + M_{(-i)} V M_{(-i)}^T \right)^{-1} M_{(-i)} X$$

$$Y_{(-i)} = \left(I - V M_{(-i)}^T \left(K_{(-i)}^{-2} + M_{(-i)} V M_{(-i)}^T \right)^{-1} M_{(-i)} \right) V$$

$$A_i = m_i Y_{(-i)} m_i^T$$

Weighting factor:

$$k_i^2 = \frac{(1 - \beta_i) r_i - D_i}{\beta_i r_i A_i}$$

end for

Step 3: convergence analysis

If the relative variations of all the weighting factors between two consecutive estimations are greater than a fixed threshold then return to step 2 otherwise stop the algorithm.

6. Fault detection and isolation

As it has been said in the introduction, data reconciliation techniques aim to enhance the quality of the measurements. It is also well known that the application of these techniques may be used for detecting and isolating sensor faults. One of the primary data analysis consists of checking the coherency of the measurements with regard to the model of the process. Usually, the vector of a priori residuals is evaluated:

$$R = M(\theta) X \tag{42}$$

In the particular case where $\theta = \theta_0 =$ constant and under the hypothesis that the measurement errors are normally distributed with zero mean and a known variance–covariance matrix V, this residual vector is also normally distributed with zero mean

and a variance-covariance matrix equal to V_{rx} [Eq. (33)]. In order to compare the components of R, let us define a normalised a priori residual vector which the ith element is defined by:

$$R_{n1}(i) = \frac{R(i)}{\sqrt{V_{rx}(i,i)}} \tag{43}$$

Each normalised residual is now normally distributed with a unity variance. A simple statistical two tailed test can, therefore, be employed to detect data inconsistency. If the absolute value of a normalised residual is greater than a fixed threshold, the corresponding equation is suspected to involve faulty measurements. Classically, one may choose this threshold to control the family-wise Type I error rate at some pre-assigned level.

If the parameters are also considered as random variables, the approximated variance–covariance matrix of the a priori residual is given by (34). Following the previous way, another normalised residuals may be defined:

$$R_{n2}(i) = \frac{R(i)}{\sqrt{V_r(i,i)}} \tag{44}$$

Due to the approximation in the calculus of variance–covariance matrix V_r, these residuals do not have any statistical properties. However, this normalisation allows them to be compared each other. As previously, the crossing of an heuristically fixed threshold by a particular normalised residual indicates the equations suspected to involve faulty measurements.

Another data analysis consists to examine the magnitude of the corrective terms caused by data reconciliation defined by:

$$E = X - \hat{X} \tag{45}$$

Let us begin by the case of a perfectly known model. In this case, from (5), we have:

$$E = VM^T(MVM^T)^{-1}MX$$

It is easy to show that this vector of corrective terms is normally distributed with zero mean and a variance–covariance matrix $V_{ex} = V - \hat{V}_x$ where $\hat{V}_x$ is defined by (11). As previously, a normalised vector of corrective terms may be defined by dividing each corrective term by its standard deviation:

$$E_{n1}(i) = \frac{E(i)}{\sqrt{V_{ex}(i,i)}} \tag{46}$$

Each of these corrective terms being normally distributed with zero mean and unity variance, if their absolute values are greater than a fixed threshold, the corresponding measurements are suspected to be faulty.

Now, as for the a priori residuals, it is also possible to evaluate this variance–covariance matrix taking into account the variance of the parameters. Let us denote θ_i, $i = 1, \cdots, p$ the components of the vector θ_0. As we only consider linear equations with regard to the parameters, the nominal matrix of constraints may always be written as:

$$M(\theta_0) = M_0 + \sum_{i=1}^{p} \theta_i M_i \tag{47}$$

The variance–covariance matrix of the estimation may be calculated by an approximation at the first order. So let us consider a small variation ΔM induced by small variations $\Delta\theta_i$ of each parameter θ_i. We can write:

$$\begin{aligned} M(\theta) &= M(\theta_0) + \Delta M \\ &= M_0 + \sum_{i=1}^{p} (\theta_i + \Delta\theta_i) M_i \end{aligned} \tag{48}$$

whence:

$$\Delta M = \sum_{i=1}^{p} \Delta\theta_i M_i \tag{49}$$

Now, let us explain $\Delta\hat{X}$ the variation of the estimation due to a variation ΔM of the constraint matrix and a variation ΔX of the measurements. For the "nominal" values, we have, from Eq. (14):

$$\hat{X} = \left(V^{-1} + M^T K^2 M\right)^{-1} V^{-1} X = A^{-1} V^{-1} X = PX \tag{50a}$$

or:

$$A\hat{X} = V^{-1}X \tag{50b}$$

In the presence of the variations ΔM and ΔX, Eq. (50b) becomes:

$$\left(V^{-1} + (M+\Delta M)^T K^2 (M+\Delta M)\right)\left(\hat{X} + \Delta\hat{X}\right) = V^{-1}(X + \Delta X) \tag{51}$$

So, limiting the expansion to the first order, the variation of the estimation may be approximated by:

$$\Delta\hat{X} = -A^{-1}\left(\Delta M^T K^2 M + M^T K^2 \Delta M\right)\hat{X} + A^{-1}V^{-1}\Delta X \tag{52}$$

or, taking into account Eq. (49) and (15a):

$$\Delta\hat{X} = \sum_{i=1}^{p} \Delta\theta_i B_i X + P\Delta X \tag{53a}$$

with:

$$B_i = -A^{-1}\left(M_i^T K^2 M + M^T K^2 M_i\right)P \tag{53b}$$

Thus, Eq. (53a) expresses the sensitivity of the estimation with regard to variations of measurements and model parameters. Consequently, it could be pointed out the most "important" parameters of the model, i.e. those having a strong influence on the estimations. On a statistical point of view, and taking into account that the variations are centered around zero, the approximated variance-covariance matrix of the estimation may be obtained from Eq. (53a):

$$\mathrm{Var}(\hat{X}) = \mathrm{E}\left(\left(\sum_{i=1}^{p}\Delta\theta_i B_i X\right)\left(\sum_{i=1}^{p}\Delta\theta_i B_i X\right)^T\right) + \mathrm{E}\left(P\Delta X(P\Delta X)^T\right) \tag{54a}$$

As the measurement and parameter errors are independent, expression (54a) reduces to:

$$\mathrm{Var}(\hat{X}) = \sum_{i=1}^{p} B_i X \mathrm{E}\left(\Delta\theta_i^2\right) X^T B_i^T + P\mathrm{E}\left(\Delta X(\Delta X)^T\right)P^T \tag{54b}$$

that is to say:

$$\mathrm{Var}(\hat{X}) = \hat{V} = \sum_{i=1}^{p} W_{ii} B_i X X^T B_i^T + PVP^T \tag{54c}$$

Using this expression, the variance–covariance matrix of the corrective terms becomes $V_e = V - \hat{V}$. Therefore, another normalised corrective terms may be defined:

$$E_{n2}(i) = \frac{E(i)}{\sqrt{V_e(i,i)}} \tag{55}$$

As for the a priori residuals, the comparison of each normalised corrective terms with a threshold helps in locating the eventually faulty measurements.

7. Numerical example

Let us consider the following set of linear (with regard to variables and parameters) equations; some of them are described with exactly known parameters, the others being characterised by uncertainties:

$$x_1^* + x_2^* - x_3^* = 0$$

$$x_3^* - x_4^* - a x_5^* = 0$$

$$x_5^* - x_6^* - b x_7^* = 0$$

$$x_4^* - c x_8^* + d x_9^* = 0$$

$$x_8^* - x_{10}^* - x_{11}^* = 0$$

$$x_{10}^* + x_{11}^* - x_{12}^* = 0$$

The measurements of the variables as well as their precision are given in Table 1.

Table 2 presents the nominal values of the parameters and their accuracy.

Table 1
Measurements and their precision

Variable	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
Measure	70	30	98	60	8	3	2	34	50	21	10	33
S.D.	3.5	1.5	4.9	3.0	0.4	0.15	0.1	1.7	1.5	1.05	0.5	1.65

The first treatment consists of checking the measurements with regard to the model. Table 3 presents the values of a priori residuals and the corresponding normalised residuals defined by Eqs. (43) and (44).

The analysis of these results clearly shows the advantages of taking into account the variance of the parameters for decision-making according to the fault detection goal. Indeed, if the parameters are considered as perfectly known, a statistical test of normal distribution of R_{n1} at the confidence level $\alpha = 0.05$ induces the conclusion that Eq. (4) is suspected to involve faulty measurements. In point of fact, taking into account the uncertainties on the parameters, one ascertains that the normalised a priori residual R_{n2} of Eq. (4) is not the greatest. From this analysis, one can conclude that all the measurements seem to be coherent with the proposed uncertain model.

After this first analysis, the data reconciliation may be envisaged. Some of the constraints are structurally exact (first, fifth and sixth equations), so the corresponding reduction factors β_i ought to be chosen equal to zero. However, in order to implement a unique procedure of calculus and taking into account that the analytical expression for calculating a weighting factor prevents the reduction factor to be null, these reduction factors are fixed to an arbitrary small value. The other constraints are uncertain. So, the reduction factors might be calculated according to (37). For example, β_4 is evaluated as follows:

$$\beta_4 = 1 - \frac{\sigma_{r_4 x}}{\sigma_{r_4}} = 1 - \frac{\left(\sigma_{x_4}^2 + c^2\sigma_{x_8}^2 + d^2\sigma_{x_9}^2\right)^{1/2}}{\left(\sigma_{x_4}^2 + c^2\sigma_{x_8}^2 + d^2\sigma_{x_9}^2 + x_8^2\sigma_c^2 + x_9^2\sigma_d^2\right)^{1/2}}$$

Applying this technique for the other constraints, one obtains the values of Table 4.

The proposed algorithm for determining the weighting factors yields the values given in Table 5.

The corresponding estimations are collected in Table 6.

In order to confirm the results obtained by the analysis of *a priori* residuals, one may compute the different corrective terms; they are presented in Table 7.

A classical analysis, considering the values of the different parameters as exact, leads to suspect, at the confidence level $\alpha = 0.05$, the measurements of x_9 and x_{10}. Indeed, the corresponding normalised

Table 2
Nominal values of the parameters and their accuracy

Parameter	a	b	c	d
Value	4	2.1	5	4
S.D.	2	0.8	2	1.6

Table 3
A priori residuals and normalised residuals

Equation	(1)	(2)	(3)	(4)	(5)	(6)
R	2	6	0.8	90	3	−2
R_{n1}	0.32	1.00	1.68	8.31	1.45	−0.99
R_{n2}	0.32	0.35	0.48	0.85	1.45	−0.99

Table 4
Reduction factors

Reduction factor	β_1	β_2	β_3	β_4	β_5	β_6
Value	10^{-5}	0.65	0.71	0.90	10^{-5}	10^{-5}

Table 5
Weighting factors

Weighting factor	k_1	k_2	k_3	k_4	k_5	k_6
Value	119.91	0.31	1.90	0.07	253.92	86.94

Table 6
Estimations

Variable	x_1	x_2	x_3	x_4	x_5	x_6
Estimation	66.48	29.35	95.83	60.29	7.91	3.05
Variable	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
Estimation	2.04	33.41	46.88	22.96	10.44	33.41

Table 7
Corrective terms

Variable	x_1	x_2	x_3	x_4	x_5	x_6
E	3.52	0.65	2.17	−0.29	0.09	−0.05
E_{n1}	1.48	0.45	0.94	−0.13	0.33	−0.32
E_{n2}	0.72	0.40	0.39	−0.05	0.15	−0.28
Variable	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
E	−0.04	0.59	3.11	−1.96	−0.44	−0.41
E_{n1}	−0.47	0.75	2.35	−2.51	−0.94	−0.52
E_{n2}	−0.40	0.49	0.73	−1.84	−0.88	−0.34

terms E_{n1} are greater than 1.96, the upper $\alpha/2$ point of the standard normal distribution. As before, a more precise analysis based on the normalised corrective terms which takes into account the variance of the parameters shows the coherency of the measurements. Therefore, it is important to note that the proposed analysis contributes a decrease in the rate of false alarms by exploiting all the available knowledge about both measurements and models.

8. Conclusion

The benefits that may be obtained by using data reconciliation techniques are demonstrated. Indeed reconciled data can be used to better monitor performances and yield for the plant and process units, aid in detecting faulty instrumentation and prioritise instrument maintenance, establish consistent sets of operation data for technical analysis and subsequent operation improvement, support the product accounting and loss control function by highlighting discrepancies between accounting and operation data. The proposed technique still enhances the benefits of applying data reconciliation techniques. Indeed, by taking into account all the available knowledge about the process model, it prevents from erroneous decisions. The involved calculus are very simple to implement and are not cumbersome. A current study aims to extend the technique to the case of non linear models.

References

[1] D. Mandel, A. Abdollahzadeh, D. Maquin, J. Ragot, Data reconciliation by inequality balance equilibration, A LMI approach. International Journal of Mineral Processing 53 (1998) 157–169.
[2] A. Abdollahzadeh, Validation de données par équilibrage de bilans, synthèse et nouvelles approches. Doctorat de l'Institut National Polytechnique de Lorraine, 27 Octobre 1997.
[3] J. Ragot, D. Maquin, F. Kratz, Observability and redundancy decomposition — application to diagnosis, in: R.J. Patton, P.M. Frank, R.N. Clark (Eds.), Advances in Fault Diagnosis for Dynamic Systems, Prentice-Hall, Control Engineering Series, Chapter 13, 1998.
[4] C.M. Crowe, Y.A. Garcia Campos, A. Hrymak, Reconciliation of flow rates by matrix projection, Part I: the linear case. AIChE Journal 29 6 (1983) 881–888.
[5] D. Maquin, G. Bloch, J. Ragot, Data reconciliation for measurements, European Journal of Diagnosis and Safety in Automation 1 2 (1991) 145–181.