SIMULTANEOUS STATE AND PARAMETER ESTIMATION. APPLICATION TO RAINFALL DATA VALIDATION.

Gilles MOUROT, Didier MAQUIN, José RAGOT

Centre de Recherche en Automatique de Nancy - CNRS UPRESA 7039
Ecole Nationale Supérieure de Géologie de Nancy
Institut National Polytechnique de Lorraine
F - 54 516 Vandoeuvre les Nancy
Email: {mourot, maquin, ragot}@ensg.u-nancy.fr

Abstract: For process control improvement, coherency of information supplied by sensors must first be ensured. Because of the presence of random and possibly gross errors, the model equations of the process are not generally satisfied. The problem of data reconciliation in order to satisfy the model constraints is considered in this article. The simultaneous presence of errors in process input and output measurements poses serious problem in the rectification of data. The proposed procedure to solve this problem involves the use of a special filter to estimate both the parameters, the states and the inputs of a process; smoothing of both estimations of the input and the output is increased by adding in the filter a variance term. Application is proposed for rainfall data validation in order to improve urban sewer network control.

Keywords: state and parameter estimation, linear dynamic model, rainfall data validation.

1. INTRODUCTION

The estimation of the state of a process is a fundamental part of modelling, monitoring and control strategies. For example, in the field of diagnosis, the success of fault detection and isolation mainly depends on the estimation of the state of the process. Generally, for diagnosis purpose, estimation has to be performed through on-line techniques either in a recursive form or not. This has been extensively studied and Kalman filter, in the stochastic case, (Karjala, 1996) or Luenberger observer, in the deterministic case, are well known approaches. Some extensions have also been considered for processes with unknown parameters; in this case, general non linear estimation involving both data reconciliation and parameter estimation has been developed (Roberston, 1996). However, in these approaches, the estimation problem is generally reduced to the state estimation, the input of the process being known.

Although our presentation is limited to linear model, the problem addressed in this article is more general than those mentioned in the previous works, since it is desired to simultaneously estimate the state, the input of the process and its parameters. In the field of process engineering, state estimation is generally seen through the classical concept of data reconciliation (Simson, 1988). Data reconciliation is mainly a physical problem: the variables of the process should obey the mass and energy conservation constraints. In the following this concept of balance constraints will be considered.

On a general point of view, the problem of state estimation of a process may be formulated as follows:

From measurements $Y_m(t)$ collected on the process, whose functioning is characterised by state variables $X(t)$, from the knowledge of the measurement $Y(t) = g(X(t))$, and from the knowledge of the model of the process $f(X(t), \theta) = 0$, where θ describes the parameters of the process, is it possible to give an estimation $\hat{X}(t)$ of the state of the process ?

Generally this problem is too complex and no analytical solution may be found. However, with some specifications on the measurement system and when considering particular descriptions of the model of the process, it is possible to establish the existence conditions of the solution and the solution itself (Bousghiri *et al.*, 1994; Karjala and Himmelblau, 1996; Liebman *et al.*, 1992). For example, it is the case when the process and the measurement system may be modelled by linear equations (with respect to the state) for which, when the parameters θ are known, the observer theory (Muske *et al.*, 1993) gives adequate solutions.

The problem under consideration here is however more general while the parameters θ are unknown. Consequently, our aim is to estimate the state of the process and simultaneously the parameters of its model. On a general point of view, this problem may be addressed as a non-linear estimation one. Cox (1964) is probably one of the first being concerned with such difficulties and has proposed an iterative solution based of the maximisation of the likelihood function of the measurement constrained by the model of the system. El Sherief (1982) has also proposed an estimation method based on the Kalman filter. These methods are also known as "Bootstrap" methods and Puthenpura (1986) has given some refinement in order to increase the robustness of the estimation. With regard to existing techniques, our contribution may be point out in the following directions. First, a complete analytic formulation of the estimation problem is presented. Second, errors affecting the measurement of both input and output of the process are taken into account. Third, data which are representative of the system, i.e. verifying all the state equations with some good properties of smoothing, are estimated.

2. GENERALISED ESTIMATION METHOD

When considering single-input-single-output system, let us note $x(t)$ the input and $y(t)$ the output, both being sampled at the same constant rate; $x_m(t)$ and $y_m(t)$ will represent the corresponding measurements. It is supposed that the true data $x(t)$ and $y(t)$ are subjected to a linear dynamic constraint:

$$y(k) = \sum_{i=1}^{n} a_i y(k-i) + \sum_{i=1}^{m} b_i x(k-i) \qquad n \geq m \tag{1}$$

From N measurements $y_m(t)$ and $x_m(t)$, the aim is to estimate the parameters a_i and b_i of the system model. Unfortunately, the measured data being subject to errors, they don't verify the model constraints; thus, one tries to simultaneously estimate the true data $y(k)$ and $x(k)$; in the following, these estimations are noted $\hat{y}(k)$ and $\hat{x}(k)$.

The principle used for the extended estimation (estimation of the parameters and the variables) is the constrained minimisation between the estimation and the measurement; here, for reason of simplicity, a quadratic criterion is suggested:

$$\Phi = \sum_{k=1}^{N} \left(\hat{y}(k) - y_m(k)\right)^2 + \sum_{k=1}^{N-1} \left(\hat{x}(k) - x_m(k)\right)^2 \tag{2}$$

Without restriction to generality, the deviation between estimation and measurement have not been weighted; however, it is easy to introduce weights, for example according to the precision of the measurement or to use more specific information about the probability density function of measurement errors. As previously said, the estimations have to satisfy the constraint:

$$\hat{y}(k) = \sum_{i=1}^{n} a_i \hat{y}(k-i) + \sum_{i=1}^{m} b_i \hat{x}(k-i) \tag{3}$$

The computation of the estimations of the variables $\hat{y}(k)$ and $\hat{x}(k)$ and the parameters a_i and b_i is achieved by minimising the criterion (2) taking into account the constraint (3) which is applied at each sampling time. On a numerical point of view, the problem is reduced to the optimisation of a quadratic criterion (with respect to state variable and parameters) subject to non-linear equality-constraints (the non-linearity resulting from the link between variables and parameters). Despite of a classical and well know formulation, the dimension of the problem (number of variables and parameters in a dynamical model) being high and noise affecting the measurements, conventional techniques for the resolution are not always powerful; consequently, we have developed and proposed an original way based on a hierarchical estimation.

3. PRACTICAL IMPLEMENTATION

In order to solve the preceding optimisation problem, the following Lagrange function is considered:

$$\begin{aligned} L = & \sum_{k=1}^{N} \left(\hat{y}(k) - y_m(k)\right)^2 + \sum_{k=1}^{N-1} \left(\hat{x}(k) - x_m(k)\right)^2 + \\ & \sum_{k=1}^{N-1} \lambda(k) \left(\hat{y}(k) - \sum_{i=1}^{n} a_i \hat{y}(k-i) - \sum_{i=1}^{m} b_i \hat{x}(k-i) \right) \end{aligned} \tag{4}$$

The complexity of the optimisation procedure is greatly simplified if a matricial presentation is used. For that purpose, let us define the following vectors of variables and parameters:

$$Z = \left(y(1) \;\; x(1) \;\; y(2) \;\; \ldots \;\; x(N-1) \;\; y(N)\right)^T \tag{5a}$$

$$\hat{Z} = \left(\hat{y}(1) \;\; \hat{x}(1) \;\; \hat{y}(2) \;\; \ldots \;\; \hat{x}(N-1) \;\; \hat{y}(N)\right)^T \tag{5b}$$

$$\theta = \left(a_1 \qquad a_n \;\; b_1 \;\; \ldots \;\; b_m\right)^T \tag{5c}$$

The constraint (3), expressed for the duration N, may be condensed under the form:

$$M(\theta)\hat{Z} = 0 \tag{6}$$

where $M \quad {}^{(N-n).(2N-1)}$ and $\theta \quad {}^{p}$ with $p = n + m$.

The Lagrangian (4) associated to the optimisation problem may be written:

$$L = \frac{1}{2}\left\| \hat{Z} - Z \right\|^2 + \lambda^T M(\theta)\hat{Z} \tag{7}$$

in which we recall the dimensions:

$$Z \quad {}^{2N-1} \qquad\qquad \hat{Z} \quad {}^{2N-1}$$

$$M \quad {}^{(N-n).(2N-1)} \qquad\qquad \lambda \quad {}^{N-n}$$

The optimality conditions are formulated (I_p being the identity matrix of dimension p):

$$\frac{\partial L}{\partial \hat{Z}} = \hat{Z} - Z + M^T(\theta)\lambda = 0 \tag{8a}$$

$$\frac{\partial L}{\partial \lambda} = M(\theta)\hat{Z} = 0 \tag{8b}$$

$$\frac{\partial L}{\partial \theta} = \left(I_p \otimes \lambda^T\right)\frac{\partial M(\theta)}{\partial \theta}\hat{Z} = 0 \tag{8c}$$

(See the involved derivative rule)[1]

Equations (8), which are non-linear, may be solved according to the following strategy. From equations (8a) and (8b), one obtains:

$$\lambda = \left(M(\theta)M^T(\theta)\right)^{-1} M(\theta)Z \tag{9a}$$

$$\hat{Z} = \left(I_{2N-1} - M^T(\theta)\left(M(\theta)M^T(\theta)\right)^{-1} M(\theta)\right)Z \tag{9b}$$

It is important to note that if $M(\theta)$ is a full row rank matrix, then the matrix $M(\theta)M^T(\theta)$ is regular. The equation (8c) being non-linear with respect to θ, this last parameter may be obtained by an iterative way. For that purpose, a Newton-Raphson algorithm is used and estimate at iteration $i+1$ is expressed:

$$\theta_{i+1} = \theta_i - \Delta\left(\frac{\partial^2 L}{\partial\theta\partial\theta^T}\right)_i^{-1}\left(\frac{\partial L}{\partial \theta}\right)_i \tag{10}$$

It is useful to write the two forms allowing the computation of the gradient of the criterion with respect to θ:

$$\frac{\partial L}{\partial \theta} = \left(I_p \otimes \lambda^T\right)\frac{\partial M(\theta)}{\partial \theta}\hat{Z} \tag{11a}$$

$$\frac{\partial L}{\partial \theta} = \left(I_p \otimes \hat{Z}^T\right)\frac{\partial M(\theta)}{\partial \theta}\lambda \tag{11b}$$

From (11a) and (11b), the second order derivative are computed:

$$\frac{\partial^2 L}{\partial\theta\partial\theta^T} = \left(I_p \otimes \lambda^T\right)M_0\frac{\partial \hat{Z}}{\partial\theta^T} + \left(I_p \otimes \hat{Z}^T\right)M_0^T\frac{\partial \lambda}{\partial \theta^T} \tag{12a}$$

with the following definition of the constant matrix M_0:

$$M_0 = \frac{\partial M(\theta)}{\partial \theta} \tag{12b}$$

The derivatives of λ and $\hat{Z}$ with respect to θ are given by:

$$\frac{\partial \hat{Z}}{\partial \theta^T} = -\left(\frac{\partial M^T}{\partial \theta^T}\left(I_p \otimes \lambda\right) + M^T(\theta)\frac{\partial \lambda}{\partial \theta^T}\right) \tag{13a}$$

$$\frac{\partial \lambda}{\partial \theta^T} = \left(M(\theta)M^T(\theta)\right)^{-1}\left(\frac{\partial M(\theta)}{\partial \theta^T}\left(I_p \otimes Z\right) - \frac{\partial\left(M(\theta)M^T(\theta)\right)}{\partial \theta^T}\left(I_p \otimes \lambda\right)\right) \tag{13b}$$

with:
$$\frac{\partial\left(M(\theta)M^T(\theta)\right)}{\partial \theta^T} = \frac{\partial M(\theta)}{\partial \theta^T}\left(I_p \otimes M^T(\theta)\right) + M(\theta)\frac{\partial M^T(\theta)}{\partial \theta^T} \tag{13c}$$

Summarising, the estimation is performed according to the following steps:

Step 1: $i = 0$
select initial values for the parameters θ_i

Step 2: compute $M(\theta_i)$
compute the state $\hat{Z}_i$ of the system (9b)

Step 3: with (14a, b, c), compute the gradient (11) and the Hessian (12a)
adjust the values of the parameters θ_{i+1}

Step 4: if the gradient of the criterion is more that a given threshold, start again at step 2, elsewhere the last estimation is considered as satisfactorily.

4. SOME SIMULATION RESULTS

The figure 1a shows the input and the output signals of a dynamical system of first order; the data are somewhat noisy. After applying the proposed estimation method to a first order model, the following parameters have been obtained:

$$\begin{cases} a_1 = 0.895 \\ b_1 = 0.202 \end{cases}$$

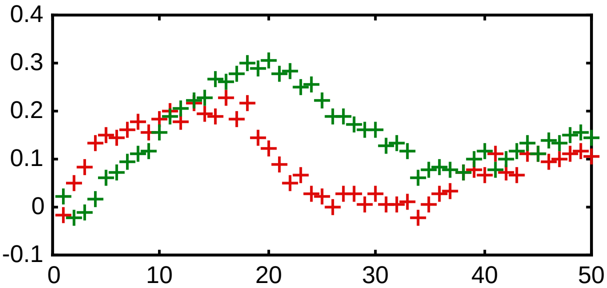

Fig. 1a. Measurement of the input and the output.

The figure 1b compares the estimated values of the input and output variables. One may appreciate the quality of the model through its ability to reconstruct the input and the output variables of the process. However, let us note that these reconstructions may be characterised by a filtering effect which seems more important for the output than for the input data. Moreover, in many situations, we have observed that, although data reconciliation is perfect, the estimation of the input is not well smoothed.

---

[1]Derivative rule of a matrix product with regard to a matrix:

$$A \in R^{n.m},\ B \in R^{m.q},\ \theta \in R^{r.s}$$

$$\frac{\partial (AB)}{\partial \theta} = \frac{\partial A}{\partial \theta}\left(I_s \otimes B\right) + \left(I_r \otimes A\right)\frac{\partial B}{\partial \theta}$$

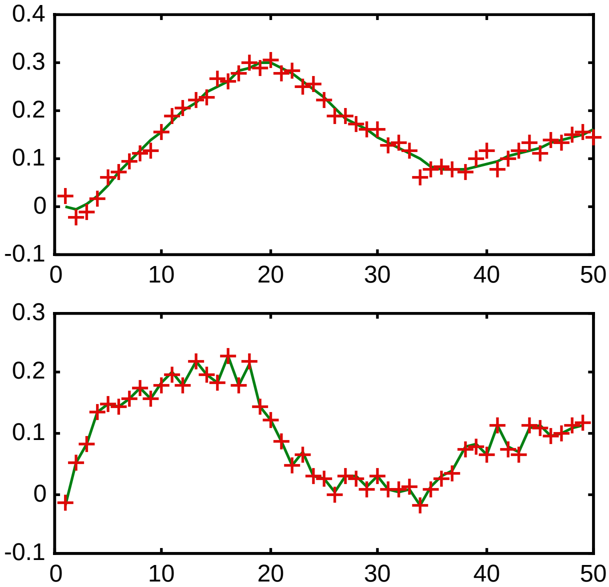

Fig. 1b. Measurement and estimation of input and output.

5. IMPROVEMENT OF THE METHOD

When measurement are corrupted by an important noise, the method (as many others) is inadequate to reconstruct an input signal with a certain degree of smoothness, although the obtained estimation perfectly verifies the state equations of the process. The figure 2 corresponds to such situation; we have used the same example as previously but involving more important noise measurements (fig. 2a).

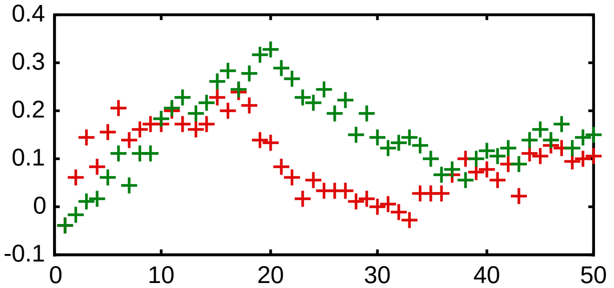

Fig. 2a. Measurement of the input and the output

It clearly appears, on figure 2b, poor filtering effect on the input signal although the output signal is correctly smoothed. It is a well known problem due to the fact that the reconstruction of an input signal generally involves an "inversion" of the transfer function of the system under consideration.

In order to reduce variation of the input estimated signal, a supplementary constraint have been added. Intuitively, by taking into account the results shown in figure 2b, the variations of the estimated input of two consecutive samples are minimised.

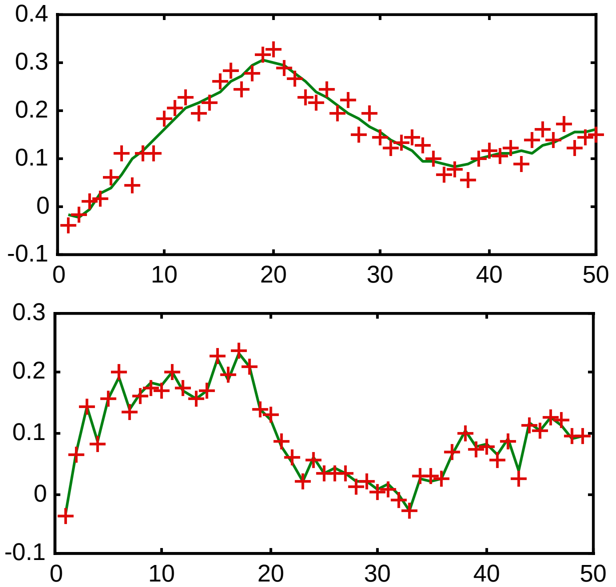

Fig. 2b. Measurement and estimation of input and output.

The quantity:

$$\phi = \sum_{k=1}^{N-1} \left(\hat{x}(k+1) - \hat{x}(k)\right)^2 \qquad (14)$$

represents the sum of squares of the variations of the estimated input. Using the following criterion thus modifies the estimation problem:

$$\begin{aligned}\Phi_m = &\sum_{k=1}^{N} \left(\hat{y}(k) - y_m(k)\right)^2 + \\ &\sum_{k=1}^{N-1} \left(\hat{x}(k) - x_m(k)\right)^2 + \alpha^2 \sum_{k=1}^{N-1} \left(\hat{x}(k+1) - \hat{x}(k)\right)^2\end{aligned} \qquad (15)$$

with respect to the constraint (3).

The parameter α^2 allows modulating the filtering effect of the estimate of the input data. Its precise choice is the fact of the user according to the particular problem to be solved.

As previously, a matricial formulation is more attractive. For that purpose, the following C matrix is defined:

$$C = \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & \dots & & & \\ & \dots & 0 & 1 & 0 & -1 & 0 & \dots & \\ \dots & & & & & & & & \\ & & \dots & 0 & 1 & 0 & -1 & 0 & \end{pmatrix}$$

that allows writing the criterion ϕ (14):

$$\phi = \left\| C\hat{Z} \right\|^2$$

The reader should modify this choice and use other filtering effects, for example, by reducing the variations of magnitude between samples $(k+2)$ and k.

The optimality conditions of the Lagrangian associated to the optimisation problem are expressed as:

$$\frac{\partial L}{\partial \hat{Z}} = \hat{Z} - Z + \alpha^2 C^T C\hat{Z} + M^T(\theta)\lambda = 0 \qquad (16a)$$

$$\frac{\partial L}{\partial \lambda} = M(\theta)\hat{Z} = 0 \qquad (16b)$$

$$\frac{\partial L}{\partial \theta} = \begin{pmatrix} I_p & \lambda^T \end{pmatrix} \frac{\partial M(\theta)}{\partial \theta} \hat{Z} = 0 \qquad (16c)$$

The resolution technique is similar as those previously given. First, the equations (16a) and (16b) are solved:

$$\lambda = (M(\theta)QM^T(\theta))^{-1} M(\theta)QZ \qquad (17a)$$

$$\hat{Z} = Q\left(I_{2N-1} - M^T(\theta)(M(\theta)QM^T(\theta))^{-1} M(\theta)Q\right)Z \qquad (17b)$$

with:

$$Q = \left(I + \alpha^2 C^T C\right)^{-1} \qquad (17c)$$

The values of θ are obtained by solving, with a recursive procedure, the equations (16c). To illustrate the performances of the proposed method, the data of the previous example are used. The weight associated to the filtering criterion is $\alpha = 2$. The figure 3 points out the filtering effect on the input data estimations and the reader could compare the results obtained without (figure 2b) and with filtering (figure 3).

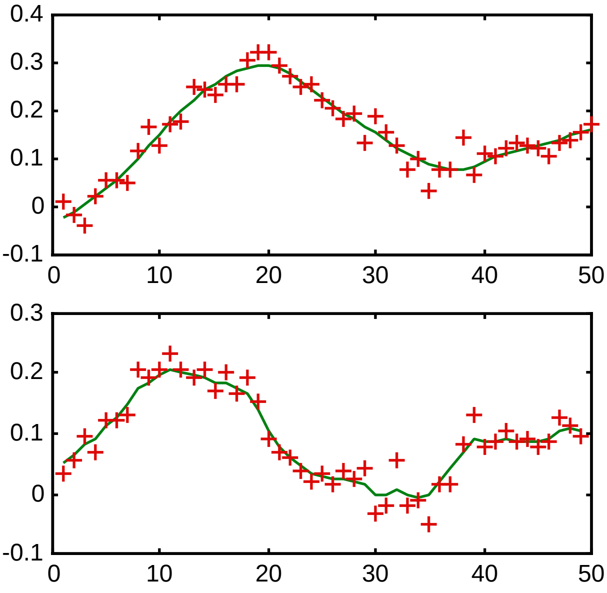

Fig. 3. Measurement and estimation of input and output.

6. RAINFALL DATA VALIDATION

The proposed technique is still under validation on a process involving rainfall data validation. The figure 4 has been drawn using the data collected during a rain event and is issued from the data bank of the Urban District of Nancy, France. The upper part of the figure relates to the precipitation (in mm/h of water) and the lower part is dedicatedto the flowrate of water in pipes (m^3/s) which constitute the input and the output of the process The parameters of a first order linear model have been estimated:

$$a = 0.6155$$
$$b = 0.0126$$

The estimations (height and flowrate) are given by the superposed curves. One can compare the measured and the estimated values and appreciate the good level of smoothing of the estimated values.

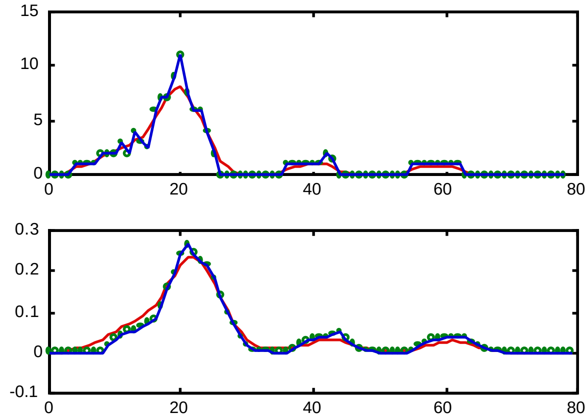

Fig. 4. Measurement and estimations of input and output

6.1 Application to sensor fault detection

The figure 5 shows the occurrence of a fault on the output sensor; this fault is a bias between samples 10 and 15. The upper part of the figure shows the measured flowrate and its estimation obtained from the model and the measurement of input and output; the lower part of the figure shows the residual between the two flowrates (estimated and measured) and the jump corresponding to the failure can be easily detected. A systematic use of this technique may be done and although the estimation is performed off line after the data collecting, it gives the user a real help for analysing the sensor behaviour.

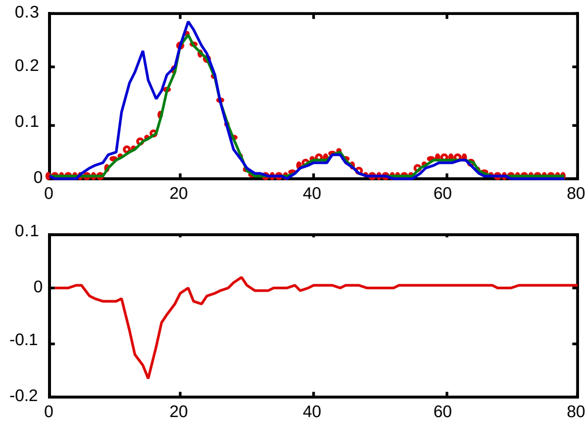

Fig. 5. Presence of a of rain gauge fault between samples 10 and 15

6.2 Extension to several events

In order to establish a more realistic model, the proposed technique may be extended when several data series are available. The problem may be formulated as follows. Each data serie is now marked by the subscript i. The vectors of the different variables are noted Z_i for the measurement and $\hat{Z}_i$ for the estimations. The dimensions of the corresponding vectors depend on data series:

$$Z_i \quad {}^{2N_i-1} \qquad\qquad \hat{Z}_i \quad {}^{2N_i-1}$$

where N_i corresponds to the number of samples in the ith campaign. However, it is important to remember that the parameters θ are constant for the different campaigns because a unique model is considered. Consequently, one must solve the stationary conditions of the Lagrangian:

$$L = \sum_{i=1}^{H} \frac{1}{2} \left\| \hat{Z}_i - Z_i \right\|^2 + \alpha^2 \left\| C\hat{Z}_i \right\|^2 + \lambda_i^T M(\theta) \hat{Z}_i \tag{18}$$

with regard to the variables $\hat{Z}_i$, λ_i, and θ. The optimality equations may be written:

$$\frac{\partial L}{\partial \hat{Z}_i} = \hat{Z}_i - Z_i + \alpha^2 C^T C \hat{Z}_i + M^T(\theta)\lambda_i = 0 \tag{19a}$$

$$\frac{\partial L}{\partial \lambda_i} = M(\theta)\hat{Z}_i = 0 \tag{19b}$$

$$\frac{\partial L}{\partial \theta} = \sum_{i=1}^{H} \left(I_p \quad \lambda_i^T \right) \frac{\partial M(\theta)}{\partial \theta} \hat{Z}_i = 0 \tag{19c}$$

These relations may be compared to those obtained in (16); so the results may be transformed into the form:

$$\lambda_i = (M(\theta) Q M^T(\theta))^{-1} M(\theta) Q Z_i \tag{20a}$$

$$\hat{Z}_i = Q\left(I_{2N_i-1} - M^T(\theta)(M(\theta) Q M^T(\theta))^{-1} M(\theta) Q \right) Z_i \tag{20b}$$

The parameters θ are estimated with a Newton-Raphson technique:

$$\theta_{j+1} = \theta_j - \Delta \left(\frac{\partial^2 L}{\partial\theta\partial\theta^T} \right)_j^{-1} \left(\frac{\partial L}{\partial \theta} \right)_j \tag{21}$$

where the components of $\frac{\partial L}{\partial \theta}$ are given by (11) and where the Hessian is computed from:

$$\frac{\partial^2 L}{\partial\theta\partial\theta^T} = \sum_{i=1}^{H} \left(\left(I_p \quad \lambda_i^T \right) M_0 \frac{\partial \hat{Z}_i}{\partial \theta^T} + \left(I_p \quad \hat{Z}_i^T \right) M_0^T \frac{\partial \lambda_i}{\partial \theta^T} \right) \tag{22}$$

It is important to notice that the analysis of several campaigns increases the robustness of the model, the parameters of the model being representative of a more important amount of data, which also increases the quality of the failure diagnosis i.e. the analysis of the residuals between the measured and estimated flowrates.

7. CONCLUSION

In this paper, an estimation technique of the parameters of a linear dynamic model when both input and output signals are corrupted by errors has been presented. It is based on a simultaneous estimation of the state and parameters. The amount of calculus is limited because a hierarchical procedure is used. The estimation of the variables is done analytically; only the estimation of the parameters requires an iterative calculus. The proposed technique could probably be extended to an on-line treatment as proposed in (Muske *et al.*, 1993; Ragot *et al.*, 1990) for state estimation.

REFERENCES

Bousghiri, S., Kratz, F. and Ragot, J. (1994). Comparison of the data reconciliation and the finite memory observer at the inverted pendulum. IFAC/IMACS Symposium Safeprocess'94.

Cox, H. (1964). On the estimation of state variables and parameters for noisy dynamic systems. *IEEE Transactions on Automatic Control*, **9**, p. 5-12.

El Sherief, H. and Sinha, N.K. (1982). Determination of the structure of a system. *IEEE Transactions on Systems, Man and Cybernetic*, **12**, p. 568-673.

Karjala, T.W. and Himmelblau, D.M. (1996). Dynamic rectification of data via recurrent neural nets and the extended Kalman filter. *AIChE Journal*, **42**, p. 2225-2239.

Liebman, M.J., Edgar, T.F. and Lasdon, L.S. (1992). Efficient data reconciliation and estimation for dynamic process using non linear programming techniques. *Computers and Chemical Engineering*, **16**, p. 963.

Muske, K.R., Rawlings, J.B. and Lee, J.H. (1993). Receding horizon recursive state estimation. Proceeding of the American Control Conference, p. 900.

Puthenpura, S. and Sinha, N.K. (1986). Robust bootstrap method for joint estimation of states and parameters of linear systems. *Journal of Dynamic Systems, Measurement and Control*, **108**, p. 255-263.

Ragot, J., Darouach, M., Maquin, D. and Bloch, G. (1990). *Validation de données par équilibrage de bilan.* Traité des nouvelles technologies, série diagnostic et maintenance, HERMES.

Robertson, D.G., Lee, J.H. and Rawlings, J.B. (1996). A moving horizon-based approach for least-squares estimation. *AIChE Journal*, **42**, p. 2209-2224.

Simson, D.E.,Everett, M.G. and Voller, V.R. (1988). Reducing the number of unknown in a constrained minimisation problem - an application to material balance. *Applied Mathematical Modelling*, **12**, p. 204-211.