Comparison of gross errors detection methods in process data

Didier Maquin, José Ragot

Comparison of Gross Errors Detection Methods in Process Data

Didier Maquin and José Ragot

Centre de Recherche en Automatique de Nancy - CNRS UA 821
BP 40 - Rue du doyen Marcel Roubault
54 501 Vandoeuvre Cedex - FRANCE

Abstract

After a short view about the fundamental problem of data reconciliation, the paper presents some majors tests (the balance residuals analysis known as method of pseudo-equations, the parity space approach, the analysis of residuals known as method of measurement test, the analysis of the sum of squares deviations) and shows their equivalence.

Introduction

In the last two decades, the problem of detecting and identifying gross errors in measurements due to biases in the measuring instruments has been well studied. Several statistical tests for this purpose have been developped, such as for example, the constraint test (CT) (Mah and al., 1976), the measurement test (MT) (Mah and Tamhane, 1982), the iterative measurement test (IMT) method, the screen combinatorial (SC) method (Serth, 1986), the dynamic measurement test (DMT) (Rosenberg, 1987), the generalized likelihood ratio (GLR) test (Narasimhan, 1987) or the maximum power (MP) test (Crowe, 1989). There are different ways to identify a large error : with a theoretical analysis of all effects leading to this error, with hardware redundancy by measuring a given process variable with different sensors, by checking the consistency of the raw data. This third alternative is selected here ; it is based on analytical redundancy by using the model equations of the process. The general procedure of error detection is classicaly divided into two main parts :

the generation of so-called residuals, which are functions of measurements that are accentuated by the errors,

the detection, the isolation and the estimation of the error.

The model of a process generally relates the state vector $x(t)$ to the input vector $u(t)$ and the output vector $y(t)$ using the matrices A, B and C. The well-known state discrete equations are written :

$$x(t+1) = A\ x(t) + B\ u(t) \qquad (1a)$$
$$y(t) = C\ x(t) \qquad (1b)$$

with $x(t)$, $y(t)$ and $u(t)$ have respective dimensions n, m and l.

After un slight rearrangement one eliminates the state vector x in order to generate the redundancy equations between the measured variables. If q is the forward shift operator ($q\ u(t) = u(t+1)$), equations (1) are rewriten :

$$y(t) = C\ (q\ I - A)^{-1}\ B\ u(t) \qquad (2)$$

Defining the observation vector :

$$z(t) = \begin{pmatrix} y(t) \\ u(t) \end{pmatrix} \qquad (3)$$

we also have :

$$M(q)\ z(t) = 0 \qquad (4)$$

Applying the equation (4) to the measurements yields a set of residuals which magnitudes are strongly related to the measurement errors devices. As it is well known examination of these residuals is achieved to detect and isolate the failure in the process or in the sensors. In process engineering practise it is a common task to use as a first data screening model that are obtained from material or energy balance equations ; they give the advantage to be structurally exact and provide a correct analysis of the residuals which are then only due to the fault measurements. The case of static balance equations has been popular in the chemical and mineralurgical engineering literature since two decades ; see by example the approach often called data reconciliation or balance equilibration.

It is clear that reconciliation and fault detection use similar techniques ; however, isolation have to precede reconciliation and the later should be successful only after isolating and removing the eroneous data. Unfortunately, as it is not always possible to detect all the bad data from direct examination, the reconciliation is performed on the raw data ; statistical tests can then be applied to analyse the obtained estimation. Hence, the methods for gross errors detection can be divided into two groups : those which apply a priori, without carrying out the full data reconciliation, by testing the redundancy equations and those which apply a posteriori on the residuals between the raw measurements and the estimations of the corresponding variables. On the application point of view both methods are used together in order to improve the robustness of the gross error detection. Based on this two approaches, many algorithms have been proposed. Most of them concern the field of chemical engineering, but some of them have been applied in the field of power

plant system and there is, in fact, no limitation on their field of application.

In the following, the processes under consideration are described by :

a model equation $\qquad M X^* = 0 \qquad$ (5a)
a measurement equation $\qquad Z = H X^* + \varepsilon \qquad$ (5b)

where X^* is the vector of the process variables, Z the measurements, H the measurements selection matrix and ε the vector of random errors characterized by its variance matrix. For reason of simplicity we focalize our attention on static system although equations (5) yields also for dynamical systems as illustrated by equation (4) ; however, the example in the last section will be given for dynamical systems.

1. Data reconciliation

The data reconciliation problem involves finding a set of adjustments to the measured data such that the adjusted values satisfy model equations (1). Following the normal distribution of the measurement errors ε, this optimization problem can be stated as :

maximize the probability density function :

$$P(Z) = \frac{1}{(2\pi)^{m/2} \, |V|^{1/2}} \exp(-\frac{1}{2} (Z-HX^*)^T V^{-1} (Z-HX^*))$$

$$(6)$$

subject to $M X^* = 0$

When the variance matrix V is known, the solution $\hat{X}$ of this problem is given by minimizing the criterion :

$$\phi = \frac{1}{2} \| Z - H \hat{X} \|_{V^{-1}}^2 \qquad (7a)$$

subject to the constraint $M \hat{X} = 0$ $\qquad$ (7b)

When the system is observable ($\text{rank}(M^T H^T) = \dim(X^*)$), a global resolution of problem (7) by the Lagrange multipliers technique leads to the unbiased estimator :

$$\hat{X} = (G^{-1} - G^{-1} M^T (M G^{-1} M^T)^{-1} M G^{-1}) \, H^T V^{-1} Z$$
$$G^{-1} = H^T V^{-1} H + M^T M \qquad (8)$$

This general expression may be simplify either if all the variables are measured (H reduces to the identity matrix) or if a preliminary extraction of the redundant part of the equation is achieved. In this later case equations (5) reduce to :

a model equation $\qquad M_r X_r^* = 0 \qquad$ (9a)
a measurement equation $\qquad Z = X_r^* + \varepsilon \qquad$ (9b)

As this system is structurally identical to system (5), we drop the subscript r and equations (8) are applied with H = I. Some matrix simplifications or a direct recalculation of the estimated may be undertaken. This leads to the classical result :

$$\hat{X} = (I - V M^T (MVM^T)^{-1} M) Z \qquad (10)$$

in which $\hat{X}$ corresponds to the estimation of the redundant variables X_r^*.

From the formula for the variance of a linear combination of random variables it can be proved that the variance of the estimated is expressed by :

$$\hat{V} = (I - V M^T (MVM^T)^{-1} M) V \qquad (11)$$

The vector E of adjustments (or residuals) and the residual criterion ϕ_R are obtained by direct substitution :

$$E = Z - \hat{X} = V M^T (MVM^T)^{-1} M Z \qquad (12)$$

$$\phi_R = \frac{1}{2} R^T (MVM^T)^{-1} R \qquad (13)$$

where $R = M Z$ $\qquad$ (14)

In order to simplify these expressions, we look for a regular matrix T which transforms the matrix M to the matrix A defined by $A = T M V^{1/2}$ with the property :

$$A A^T = I \qquad (15)$$

It can be found that this transformation yields to the orthogonalisation of the rows of the matrix M ; the classical Gram-Schmidt procedure may be applied for that purpose.

With $M = T^{-1} A V^{-1/2}$ reported in (10), (11), (12) and (13), we obtain the estimated, the variance, the residuals, and the criterion :

$$\hat{X} = (I - V^{1/2} A^T A V^{-1/2}) Z \qquad (16)$$

$$\hat{V} = (I - V A^T A) V \qquad (17)$$

$$E = V^{1/2} A^T A V^{-1/2} Z \qquad (18)$$

$$\phi_R = \frac{1}{2} Z^T V^{-1/2} A^T A V^{-1/2} Z \qquad (19)$$

The formula (10, 11, 12, 13) or equivalently (16, 17, 18, 19) completely define the adjusted values, their variance and the residual criterion. Unfortunately, if one or more gross error (due by example to a corrupted measurement, a leak in a physical process or a fluctuation of the process) is present in the measurements, the reconciliation of data will be strongly affected ; moreover, the statistical hypothesis of purely random errors is not valid. It is then a necessity to detect and isolate these bad data. We present now the different methods in use to solve this problem.

2. Gross errors detection algorithm

Measurement adjustments analysis

The adjustment vector E can be tested to detect a gross error by analysis the magnitude of its different terms. From equation (12) the variance matrix of the adjustment terms may easily be computed :

$$V_E = V\, M^T\, (MVM^T)^{-1} M\, V \qquad (20)$$

so that elements $E(i)$ of E could be tested against the unit normal variate :

$$E_N(i) = \frac{E(i)}{\sqrt{V_E(i,i)}} \quad \text{for } i = 1, ..., n \qquad (21)$$

Instead of E, a linear combination of E ($T = S\,E$) may also be used. Moreover the choice $S = V^{-1}$ gives the maximum power test since it will be expected to exceed, on absolute value, any other statistic for that error. Thus it has the greatest probability to reject the null hypothesis that is no gross error in the data.

Instead of $E_N(i)$ we may also use the more compact definition :

$$E_N = \text{diag}(V_E)^{-1/2}\, E \qquad (22)$$

Using the matrix A, (22) reduces to :

$$E_N = \text{diag}(A^T A)^{-1/2}\, A^T A V^{-1/2}\, Z \qquad (23)$$

From these expression, it would be demonstrate later that when a gross error exists in the data, the greatest term $E_N(i)$ of the vector E_N corresponds to the number of the corrupted data.

The generalized likelihood ratio method

This method has been introduced by several authors, among which Willsky and Jones (1976). When a gross error of magnitude δ is present in the ith measurement, we can write :

$$Z = X + \varepsilon + \delta\, e_i \qquad (24)$$

where e_i is the ith vector of the identity matrix.

If no gross error are present (null hypothesis H_0) R has the following properties :

$$\text{Exp}(R) = 0 \qquad (25a)$$
$$\text{Var}(R) = MVM^T \qquad (25b)$$

If a gross error is present in measurement i (alternative hypothesis H_1), we can show that :

$$\text{Exp}(R) = M\, \delta\, e_i = \delta\, m_i \qquad (26a)$$
$$\text{Var}(R) = MVM^T \qquad (26b)$$

where m_i is the ith column of M.

In order to test the hypothesis, we make use of the likelihood ratio test defined in our case by :

$$r_i = \frac{\text{Pr}(R \mid H_1)}{\text{Pr}(R \mid H_0)} \qquad (27)$$

Using the normal probability density function for R, and with $V_R = MVM^T$, we obtain :

$$r_i = \frac{\exp(-\frac{1}{2}(R - \delta\, m_i)^T V_R^{-1} (R - \delta\, m_i))}{\exp(-\frac{1}{2} R^T V_R^{-1} R)} \qquad (28)$$

Since the log function is monotonic, instead of (28) we use :

$$s_i = 2\,\text{Log}(r_i) = R^T V_R^{-1} R - (R - \delta\, m_i)^T V_R^{-1}(R - \delta\, m_i) \qquad (29)$$

The computation proceeds in two steps. First, for any vector m_i, we compute the estimate of δ ; direct search for the maximum of s_i yields :

$$\delta = (m_i^T V_R^{-1} m_i)^{-1} (m_i^T V_R^{-1} R) \qquad (30)$$

Then, substituting this value in (29), we obtain the corresponding value of s_i :

$$s_i = \frac{(m_i^T V_R^{-1} R)^2}{m_i^T V_R^{-1} m_i} \qquad (31)$$

This calculation has to be performed for every vector m_i and the supremum test is :

$$s = \sup(s_i) \quad i = 1, ..., n \qquad (32)$$

Let m_i the vector that leads to the optimum in equation (32). The test statistic s is compared with a prespecified threshold ; if s is greater than this threshold then a gross error has been detected and its magnitude is estimated with (30).

Residual criterion analysis

Criterion function test
The gross errors (in the linear case) can be identified by examining the objective function. The quantity $\phi_R = E^T V^{-1} E$, with $E = \hat{X} - Z$, has a chi-square distribution with the number of degrees of freedom equal to the rank of A. Thus the imbalances of the equations can be globally tested against tabulated values of chi2. If all gross errors in measurement have been correctly deleted, the function ϕ_r will be below the threshold for the chi-square with the appropriate confidence level. Unfortunately, it does not proove, if the chi-square is verify, that there are no gross errors in the measurement set ; a gross error may exist among a large set of measurements.

A recursive algorithm has also been proposed which results from the deletion of suspect measurements ; some algebraic manipulation avoids the inversion of large matrices which appear in the computation of the projection matrix. It is shown that if the deletion of a single measurement decreases the objective function (in respect to a statistical test), this measurement corresponds to a gross error.

Criterion sensitivity in respect to the measurements

Let us consider the expression (13) of the residual criterion in order to examine the effects on a modification ΔZ of the measurements Z. As the criterion is a quadratic form in terms of Z, we can directly derive the expression of the sensitivity vector :

$$S = \frac{\partial \phi_R}{\partial Z} = M^T (MVM^T)^{-1} M Z \qquad (33)$$

The matrix variance of the sensitivity vector is :

$$Var(S) = M^T (MVM^T)^{-1} M \qquad (34)$$

and the normalized sensitivity vector S_N :

$$S_N = diag(Var(S))^{-1/2} S \qquad (35)$$

which can be also expressed with the matrix A :

$$S_N = diag(A^TA)^{-1/2} A^TA V^{-1/2} Z \qquad (36)$$

The sensitivity vector may be analyzed in respect to the error measurement. Using equation (24) and neglecting the random term ε gives :

$$S_N = \delta \, diag(A^TA)^{-1/2} A^TA V^{-1/2} e_i \qquad (37)$$

If α_{ij} are used for the terms of the matrix A^TA, then the kth term of S_N is :

$$S_N(k) = \delta \, \alpha_{kk}^{-1/2} \, \alpha_{ki} \, v_{ii}^{-1/2} \qquad (38)$$

Two of the terms of S_N are compared with the ratio :

$$\frac{S_N(k)}{S_N(i)} = \frac{\alpha_{ki}}{\alpha_{kk}^{-1/2} \, \alpha_{ii}^{-1/2}} \qquad (39)$$

Using the definition of the terms α_{ki} and remembering the triangulary inequality, it is then clear that this latter ratio is always less than one. The greatest term of S_N is the one which is corrupted by the gross error.

Sequential error detection by measure suppression

A difficulty with this global test is that while it indicates well the presence of gross errors it is not able to identify the source of these errors. The use of a sequential procedure allows one to locate the streams which contain a gross error. For the set of all process measurements one first calculates the global test ϕ_R ; if an error is indicated by the test, all measurements are considered as suspect candidates. Then, the measurements are "deleted" sequentially from the process (in groups of size 1, 2, ...). After each deletion the global test is again applied. In this approach we wish to assess the effect of deleting a particular set of measurement on the objective function and on the estimations. Moreover, it is possible to have the same approach as the one developed in the case of multiple-observer for state reconstruction (Frank, 1989) by comparing together the different estimations obtained after each deletion. It is also possible to consider suspect measurements by assigning them an infinite variance. The corresponding variation of the criterion ϕ is then used to detect the possible gross errors. By isolating the measurement x_2, for which the variance will be later modified, let us consider the following partitioning of the matrices :

$$M = [M_1 \; m_2] \quad and \quad X = [X_1 \; x_2] \qquad (40)$$

In the same time, let us consider a modification Δv_2 of the variance of this measurement. Then the whole variance matrix is written :

$$V + \Delta V = \begin{pmatrix} V_1 & 0 \\ 0 & v_2 + \Delta v_2 \end{pmatrix} \qquad (41)$$

The residual criterion (13) is then modified :

$$\phi_R + \Delta \phi_R = \frac{1}{2} \, R^T M^T (M(V + \Delta V)M^T)^{-1} R \qquad (42)$$

from which, when Δv_2 is infinite, the following variation can be deduced :

$$\Delta \phi_R = - \frac{R^T K m_2 m_2^T K R}{m_2^T K m_2} \qquad (43)$$

with $K = (MVM^T)^{-1}$

Equation (43) gives a simple expression of the reduction in the objective function when deleting a single measurement. Then, aside from vector-matrix multiplications, the only computational effort needed is the calculus of K which is done once and once only whatever is the suspect variable.

Equation residuals analysis

Generally, the measurement vector Z does not satisfy the constraint equations. With the previous hypothesis of a gaussian distribution of the measurement errors, one shows that the imbalance residuals vector R follows a normal distribution with zero mean and covariance V_R :

$$V_R = M V M^T \qquad (44)$$

In order to compare the elements of the R vector, let us define a normalized imbalanced vector R_N whose element $R_N(i)$ is defined by :

$$R_N(i) = \frac{R(i)}{\sqrt{V_R(i,i)}} \quad for \; i = 1, ..., n \qquad (45)$$

Each element $R_N(i)$ has a normal distribution with a zero mean and unity variance. Then, a simple statistical test criterion of data inconsistency can be used. From a cumulative normal distribution table the probability of $R_N(i)$ being, for example, in the interval of 1.96 to 1.96 is read to be 0.95. Therefore, when $| R_N(i) | > 1.96$, we might say that the inconsistency is significant with a probability of 0.95. This denotes that equation i is a bad equation. If we assume the presence of only one gross

error which affects only one stream of the process, it has been established that the equation which contains this stream has the bigger residual. Then, a simple examination of the elements of R shows the suspect stream.

When several gross errors are present, their location is more complicated. It has been proposed (Mah, 1982) to apply the preceding test to each equation and also to the aggregates of two or more equations, which are called pseudoequations.

Parity space approach

In the absence of gross errors, the measurements depend on the true values following the linear relation :

$$Z = HX^* + \varepsilon \qquad (46)$$

where Z is the (v.1) measurement vector, H the (v.m) measurement matrix, X^* the (m.1) vector of true values and ε the (v.1) measurement errors vector with variance V.

The parity vector is related to the measurement vector Z through a projection matrix W of dimension n.v (n=v-m) :

$$P = W \, V^{-1/2} \, Z \qquad (47)$$

where the matrix W has the following properties :

$$\begin{aligned} W H &= 0 \\ W W^T &= 0 \\ W^T W &= I - V^{-1/2} H (H^T V^{-1/2} H)^{-1} H^T V^{-1/2} \end{aligned} \qquad (48)$$

Equations of parity show that for normal functioning, the magnitude of parity vector is small (presence of measurement noise). If a failure occurs in only one of the sensors, then the parity vector may grow in a fixed direction associated with the failed sensor. Moreover the components of the parity vector have the same probability distribution as the measurement errors which are independent gaussian with a zero mean value. The variance-covariance matrix of the parity vector P is given by :

$$V_P = I \qquad (49)$$

As the variable $c^2 = P^T V_P^{-1} P$ is the sum of the square of (v-m) normally distributed variables, it has a chi-square probability distribution with (v-m) degrees of freedom and may be compared to the threshold $c^2_{1-\alpha}$ where $c^2_{1-\alpha}$ is the value of chi-square at a confidence level α. Once the detection of gross errors has been made, we try to locate them. For each column W_j of the projection matrix W, we compute the projection of the parity vector which is given by :

$$P_j = \frac{W_j^T \, P}{\| W_j \|} \qquad (50)$$

More globally, the projection vector is then given by :

$$\text{Proj}(P/W) = \text{diag}(W^T W)^{-1/2} W^T P \qquad (51)$$

When the kth sensor is faulty, then the parity vector P follows the direction of the kth column of W which is the greatest projection of P on the axis. Then after the detection of several failed sensors, we locate those with the greatest projection. Next, we delete the suspect sensor and calculate the detection test after the deletion of each sensor. We stop the procedure when the magnitude of the parity vector P corresponding to the remaining sensors no longer satisfies the detection test.

3. Comparisons

Normalized correctives terms and parity vector

Let us remember the expression of the corrective terms :

$$E_N = \text{diag}(A^T A)^{-1/2} A^T A V^{-1/2} Z \qquad (52a)$$
$$E_N = \text{diag}(A^T A)^{-1/2} A^T T R \qquad (52b)$$

We note a strong connexion between the expression of the parity vector projection onto the fault directions (50) and those of the corrective terms (52). Therefore, we examine the equivalence of the playing role of the matrices A and W. In order to define precisely this equivalence, we first proceed with the elimination of the constraint $M X^* = 0$:

We extract from M its greater regular part M_1 :

$$M = [\, M_1 \quad M_2 \,] \qquad (53)$$

The vector of true data X^* may be decomposed following this partitionning :

$$X^* = \begin{pmatrix} X_1^* \\ X_2^* \end{pmatrix} \qquad (54)$$

The true data verify the constraint :

$$M_1 X_1^* + M_2 X_2^* = 0 \qquad (55)$$

As M_1 is a regular matrix, X_1^* may be expressed :

$$X_1^* = - M_1^{-1} M_2 X_2^* \qquad (56)$$

We then obtain for the whole vector :

$$X^* = C X_2^* \qquad (57)$$

where :

$$C = \begin{pmatrix} - M_1^{-1} M_2 \\ I \end{pmatrix} \text{ with dimension } v.(v-n) \qquad (58)$$

Then, the measurement equation take a form which looks like (5b) :

$$X = C X_2^* + \varepsilon \qquad (59)$$

The parity vector P is then defined by :

$$P = W \, V^{-1/2} \, X \qquad (60)$$

According to the preceding formulation, the estimation problem is reduced to find the minimum of the criterion ϕ, in respect to $\hat{X}_2$:

$$\phi = \frac{1}{2} \| C \hat{X}_2 - X \|^2 \qquad (61)$$

the solution yields :

$$\hat{X}_2 = (C^T V^{-1} C)^{-1} C^T V^{-1/2} X \qquad (62)$$

and for the complete vector :

$$\hat{X} = C \hat{X}_2 \qquad (63)$$

$$\hat{X} = C (C^T V^{-1} C)^{-1} C^T V^{-1/2} X \qquad (64)$$

Moreover, if we compare the estimations (16) and (64) :

$$A^T A = I - V^{-1/2} C (C^T V^{-1} C)^{-1} C^T V^{-1/2} \qquad (65)$$

And comparing (48) et (65) :

$$W^T W = A^T A \qquad (66)$$

On the other hand, we simply verify that :

$$W^T P = A^T T R \qquad (67)$$

The equations (66) and (67) demonstrate the identity of the effect of the matrices W and A and also the identity between the projections of the parity vector (51) and the vector of the normalized corrective terms (52b). The comparison would be achieved when its can be proved that A follows the two supplementary properties conditions given by Potter and Suman. These conditions may be easily fulfilled by a judicious choice of the method used for the orthonormalisation of A.

Orthogonal matrix and sensitivity of the criterion

Sensitivity of the criterion related to a modification of the measurement variances can be well studied when an orthogonal matrix is used. From equation (42), its can be obtained the criterion variation :

$$\Delta\phi_R = - \frac{(R^T K m_2)^2}{\Delta v_2^{-1} + m_2^T K m_2} \qquad (68)$$

With the change of matrix :

$$A = T M V^{1/2} \qquad (69)$$

we also have :

$$M = T^{-1} A V^{-1/2} \qquad (70)$$

The new incidence matrix A is partitionned :

$$A = [A_1 \ a_2] \qquad (71)$$

which terms are defined by :

$$M_1 = T^{-1} A_1 V_2^{-1/2} \qquad (72)$$

$$m_2 = T^{-1} a_2 v_2^{-1/2} \qquad (73)$$

Reporting (73) in (68) then gives :

$$\Delta\phi_R = - \frac{(R^T T^T a_2)^2}{v_2 \Delta v_2^{-1} + a_2^T a_2} \qquad (74)$$

In the case of an infinite variance, (74) reduces to :

$$\Delta\phi_R = - \frac{(R^T T^T a_2)^2}{a_2^T a_2} \qquad (75)$$

In this expression let us remember that T is a transformation matrix (eq 69), R is the vector of equation residuals and a_2 states for the column of the incidence matrix of the removed data corresponding to an infinite variance. It is then easy to give the influence of the removing of a each measurement by considering that a_2 is one of the columns of the matrix A. One then obtain the vector of the variations of the criterion when each measurement has been successively removed :

$$\Delta\phi_R = - \text{diag}(A^T A)^{-1} A^T T R .* A^T T R \qquad (76)$$

where the operator .* is used to calculate the product of two vectors element by element.

When comparing (52) and (76) we note that the variation of the criterion due to a measurement deletion corresponds, except for the sign, to the square of the normalized corrective term corresponding to this stream before the measurement has been removed.

GLR and comparison

Let us return to expression (31) of the log of the likelihood function. By substituting R by its definition (14) and H by its definition, we obtain :

$$s_i = \frac{(e_i^T M^T K M Z)^2}{e_i^T M^T K M e_i} \qquad (77)$$

It is clear that this expression is identical, except for the sign, to the sensitivity of the criterion obtained in (43).

4. Numerical results

For the class of dynamical systems described by eq (1) the above described results when apply directly to the filtering of both the state and the input.

$$X_{k+1} = A X_k + B U_k \qquad (78a)$$
$$Y_k = C X_k \qquad (78b)$$

On a window of length N, the constraint equations (78a) are condensed :

$$M \hat{X} = 0 \qquad (79)$$

with :

$$M = \begin{pmatrix} A & B & -I & . & . & . & . & . \\ . & . & A & B & -I & . & . & . \\ . & . & . & . & A & B & -I & . \\ . & . & . & . & . & . & . & . \end{pmatrix} \qquad (80)$$

where, for simplification, the "." stand for zero blocks, and with :

$$X = (X_0 \; U_0 \; X_1 \; U_1 \; X_2 \; U_2 \; ... \; U_N)^T \qquad (81)$$

The measurement equation (78b) is writen :

$$Z = H X \qquad (82)$$

As the inputs and only a part of the state (equation 1b) are considered to be measured, we have the following definition of the selection matrix H :

$$H = \begin{pmatrix} C & . & . & . & . & . & . & . \\ . & I & . & . & . & . & . & . \\ . & . & C & . & . & . & . & . \\ . & . & . & I & . & . \end{pmatrix} \qquad (83)$$

When considering the maximum likelihood, the evaluation balance problem consists in retrieving the minimum value with respect to $\hat{X}$ of the function :

$$\phi \; = \; \frac{1}{2} \; \| Z \; - \; H \, \hat{X} \|^2_{V^{-1}} \qquad (84)$$

under the constraint (79) where $\hat{X}$ represents the vector of the estimated values and V stands for the variance-covariance matrix (the coefficients of which being all known) of the error measurements.

The solution is given by eqn (8) where it can easily be proved that G is a regular matrix if the matrix M of constraints is of full row rank. The formula (8) can also be used only to update the terms "situated" in a window of a width corresponding to the significant terms of the projection matrix. This window width corresponds to the memory of the procedure ; it is related to the matrix of the state equation and to the noise level affecting the measurements of both the state X and the control U. As previously explained in the first section, the corrective terms are then computed ; their analysis on a sliding window allows the detection of abnormal changes due to faulty measurements.

A third order system has been simulated ; the corresponding state equations are defined by :

$$A = \begin{pmatrix} 0.77 & 0.19 & 0.00 \\ -0.39 & 0.58 & 0.30 \\ -0.60 & -0.45 & 0.86 \end{pmatrix} \qquad B = \begin{pmatrix} 0.053 \\ 0.389 \\ 0.126 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (85)$$

In this simulation example, several tests have been performed in order to improve the detection algorithm. The figure 1 shows the states and the input measurements; differents measurements faults have been incorporated (for state 1 between time 22 and 242, for state 2 between time 111 and 131, for input between time 333 and 343). The figure 2 gives the evolutions of the corrective terms by using all the state measurements accordingly to the definition of C. A detection of jumps using, by example, the Page-Hinkley test enables the detection of the faults.
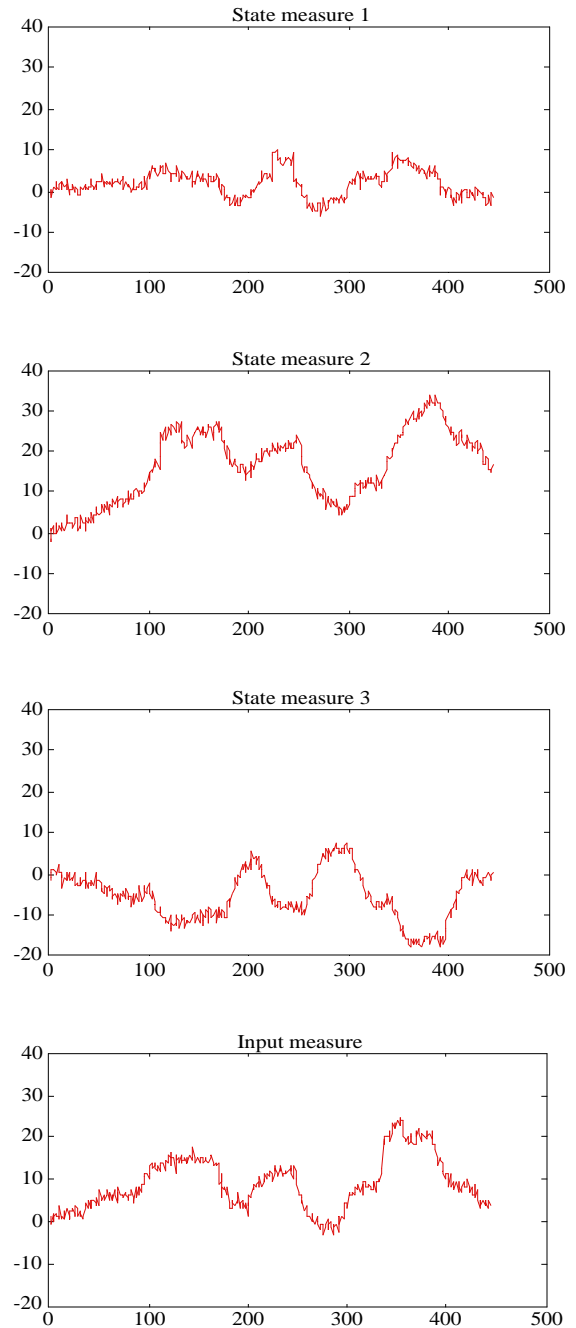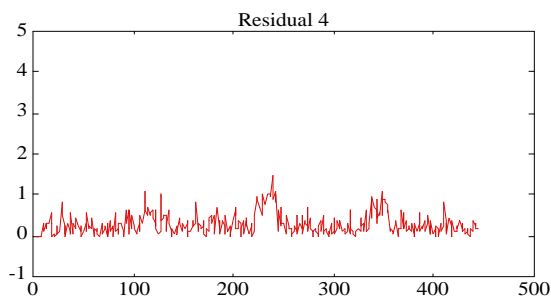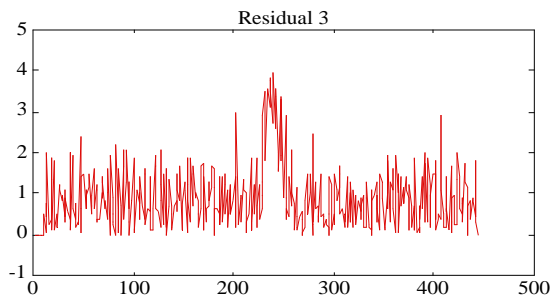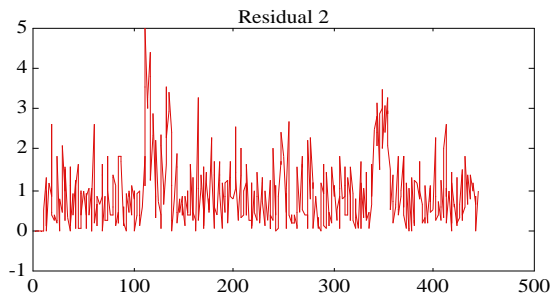

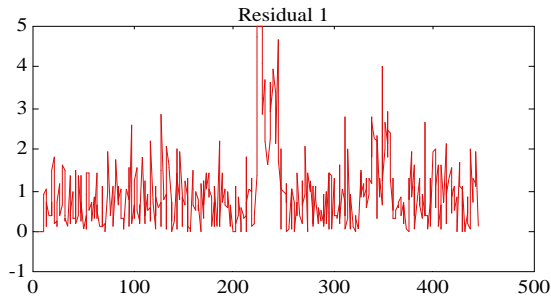Figure 1 : states and input measurements

Figure 2 : residuals

In conclusion we have prooved the equivalence of three tests commonly used for the gross errors detection purpose : parity vector, normalized corrective terms, GLR test, variation of the residual criterion after measurement deletion.

References

[CLAR 78] R.N. CLARK, "Instrument fault detection", IEEE Trans. on AES, vol. 14, n° 3, p. 456-465, 1978.

[CROW 89] C.M. CROWE, "Test of maximum power for detection of gross errors in process constraints", AIChE J., vol. 35, n° 5, p. 869-872, 1989.

[FRAN 89] P.M. FRANK, J. WUNNENBERG, "Fault diagnosis based on state estimation with reduced parameter sensitivity". Fault diagnosis in dynamic systems : theory and application. Prentice Hall, 1989.

[FRAN 90] P.M FRANK, "Fault diagnosis in dynamical systems using analytical and knowledge based redundancy - a survey and some new results", Automatica, vol. 26, n° 3, 1990.

[MAH 76] R.S.H. MAH, G.M. STANLEY, D.M. DOWNING, "Reconciliation and rectification of process flow and inventory data", I E C Proc. Des. Dev., vol. 15, n° 1, p. 175-183, 1976.

[MAH 82] R.S.H. MAH, A.C TAMHANE, "Detection of gross errors in process data", AIChE J., vol. 28, n° 5, p. 828-830, 1982.

[NARA 87] S. NARASIMHAN, R.S.H. MAH, "Generalized likelihood ratio method for gross error identification", AIChE J., vol. 33, n° 9, p. 1514-1521, 1987.

[PATT 89] R.J. PATTON, P.M. FRANK, R.N. CLARK. Fault diagnosis in dynamic systems : theory and application. Prentice Hall, 1989.

[POTT 77] J.E. POTTER, M.C. SUMAN, "Thresholdless redundancy management with arrays of skewed instruments". Electronic Flight Control systems, Agardograph 224, p. 2115-21, 1977.

[RAGO 90] J. RAGOT, M. DAROUACH, D. MAQUIN, G. BLOCH, Validation de données et diagnostic. Traité des nouvelles technologies, Hermès, 1990.

[ROSE 87] J. ROSENBERG, R.S. MAH, C. IORDACHE, "Evaluation of schemes for detecting and identifying gross errors in process data", I E C Research, vol. 26, n° 3, p. 555-564, 1987.

[SERT 86] J.E. SERTH, W.A. HEENAN, "Gross error detection and data reconciliation in steam-metering system". AIChE J., vol. 32, n° 5, p. 733-742, 1986.

[TAMH 82] A.C. TAMHANE, "A note on the use of residuals for detecting an outlier in linear regression", Biometrika, vol. 69, n° 2, p. 488-489, 1982.

[WILL 76] A.S. WILLSKY, H.L. JONES, "A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems". IEEE Trans. on Automatic Control, vol. 21, p. 108-112, 1976.