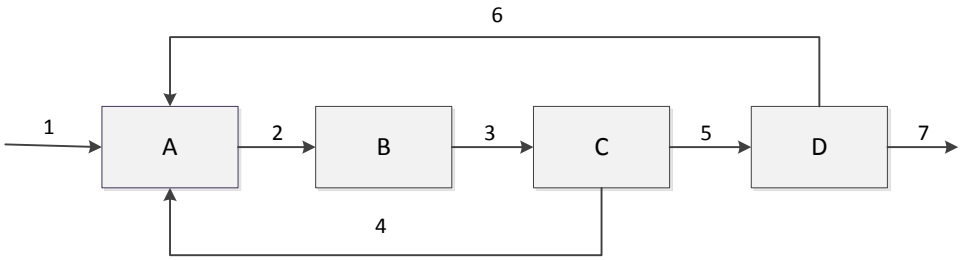
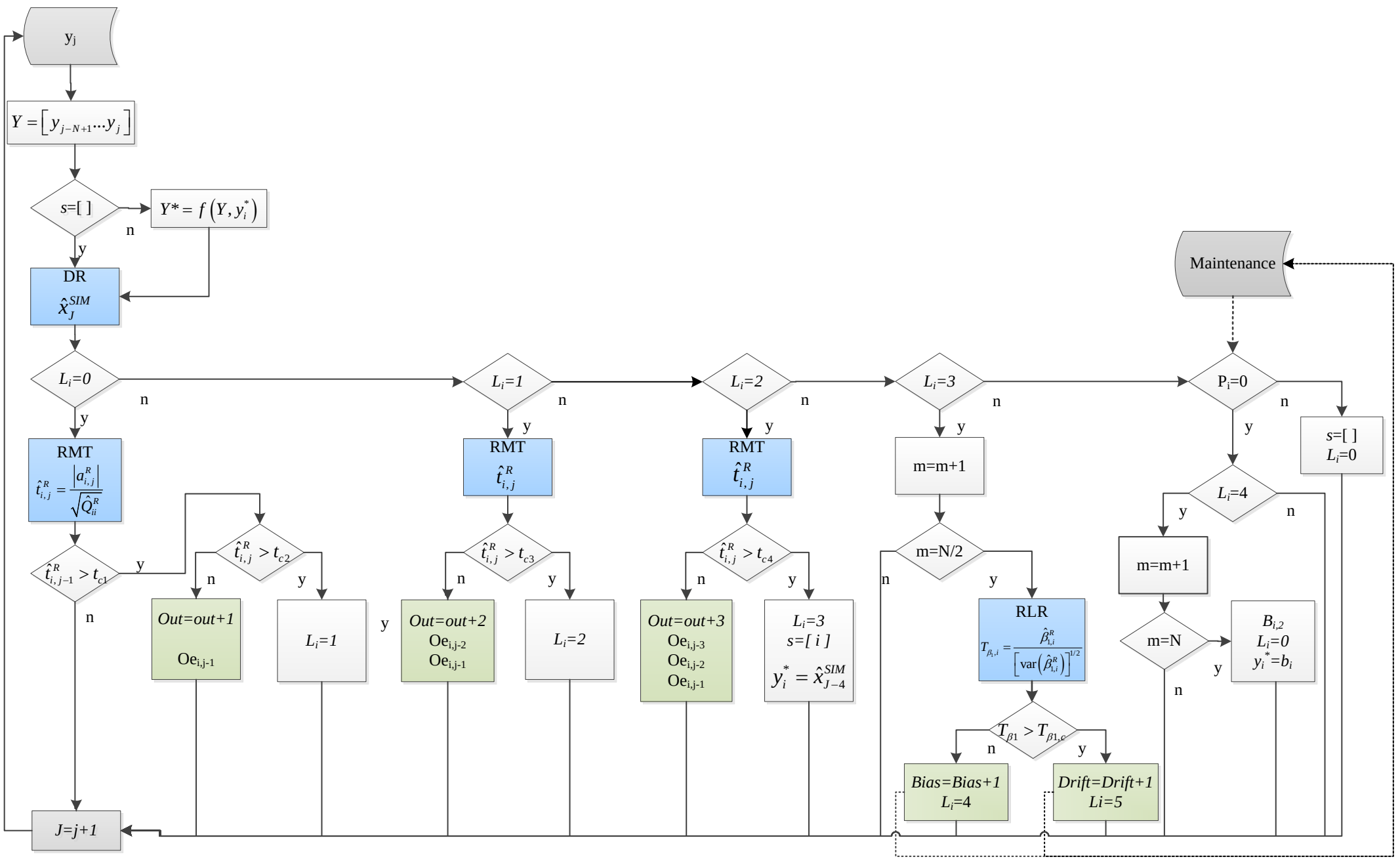
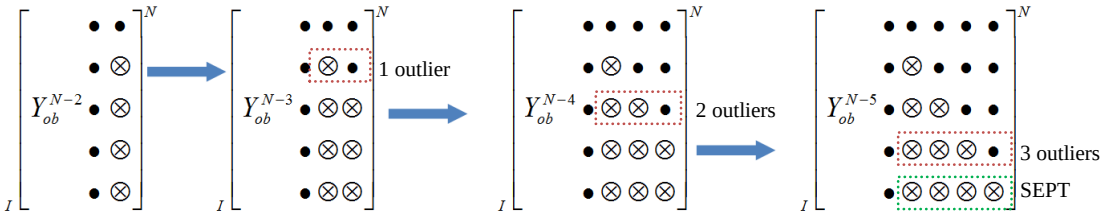
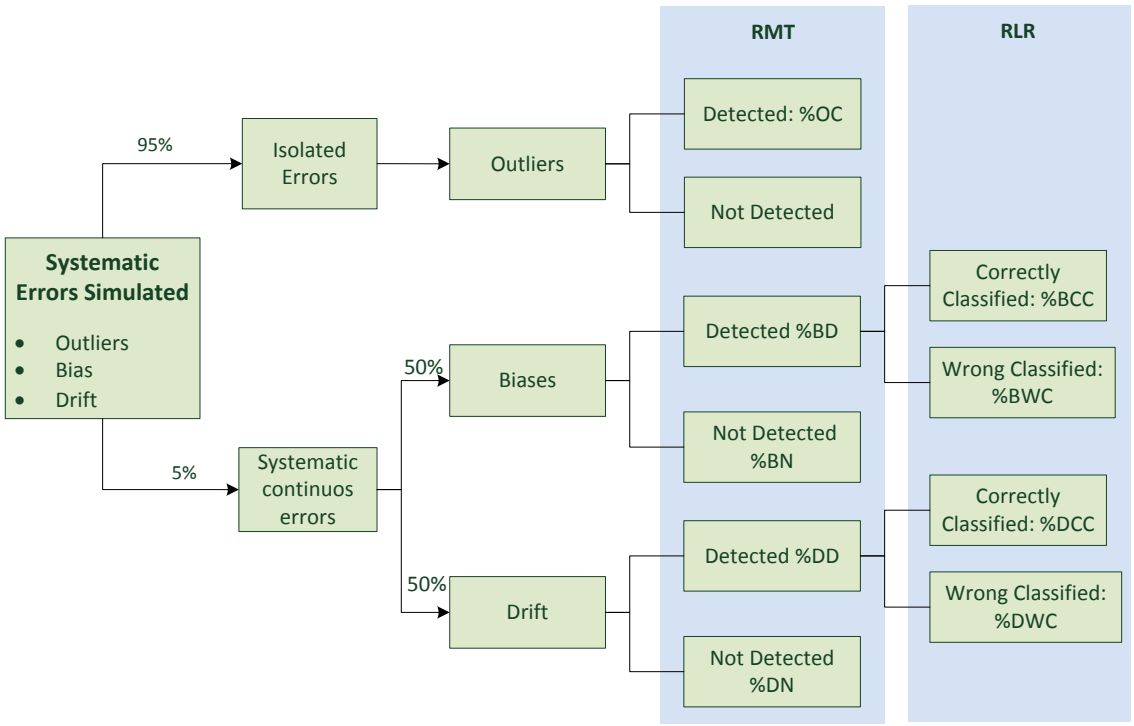
CLASSIFICATION OF SYSTEMATIC MEASUREMENT ERRORS WITHIN THE FRAMEWORK OF ROBUST DATA RECONCILIATION

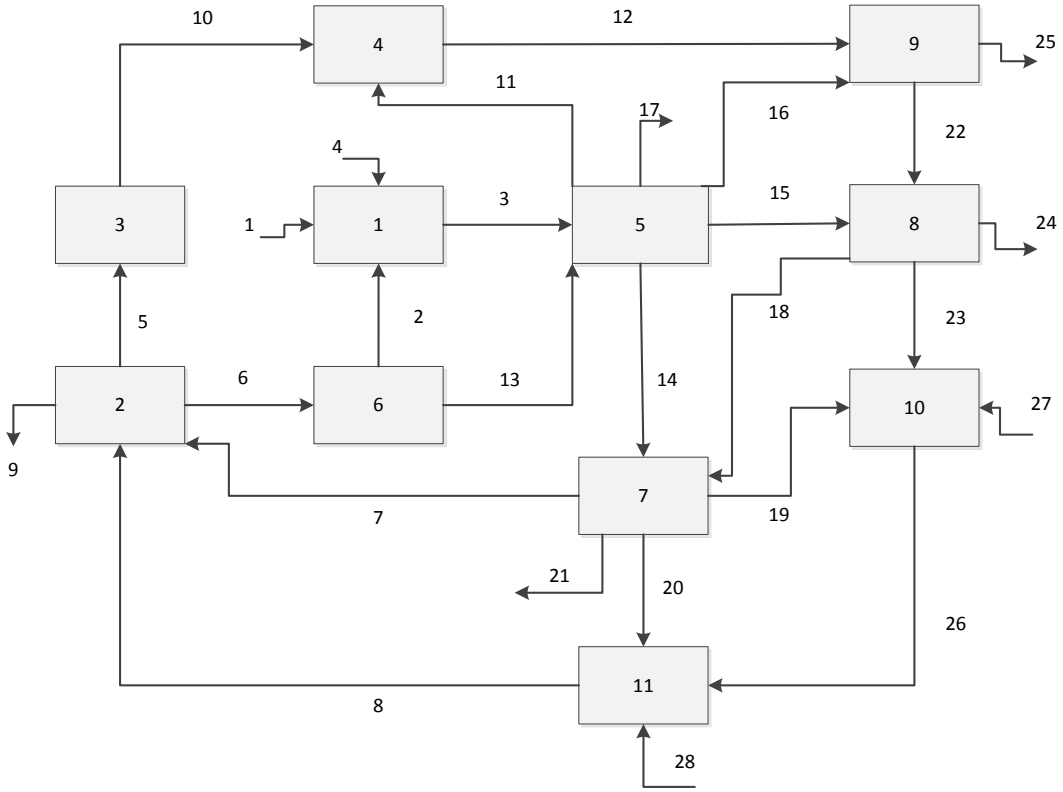
Journal:	*Industrial & Engineering Chemistry Research*
Manuscript ID	ie-2017-00726n.R2
Manuscript Type:	Article
Date Submitted by the Author:	12-Jul-2017
Complete List of Authors:	Llanos, Claudia; Planta Piloto de Ingenieria Quimica, Departamento IQ – Universidad Nacional del Sur (UNS) Bahía Blanca, Argentina Sánchez, Mabel; Planta Piloto de Ingenieria Quimica, Departamento IQ – Universidad Nacional del Sur (UNS) Bahía Blanca, Argentina Maronna, Ricardo; Universidad Nacional de La Plata, Matemática

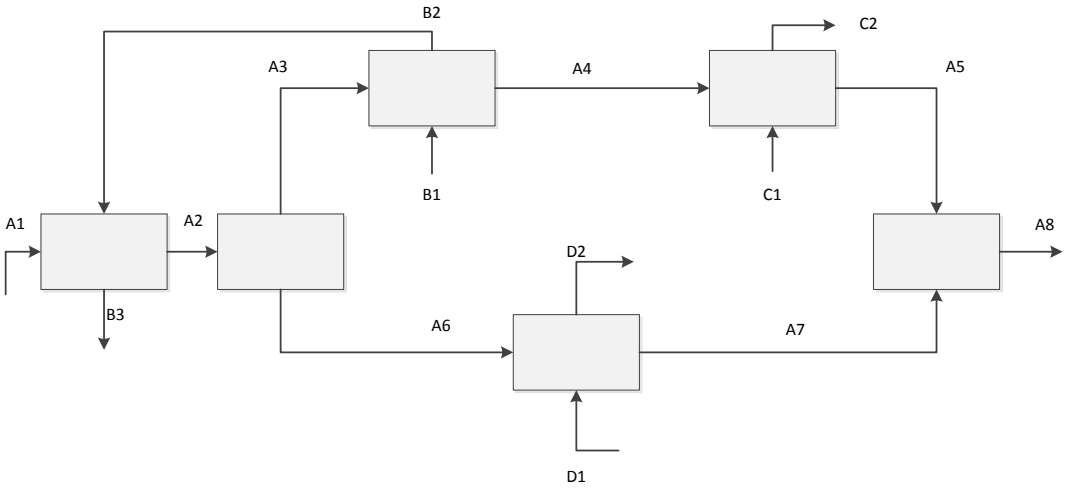

6
1
A
2
B
3
C
5
D
7
4

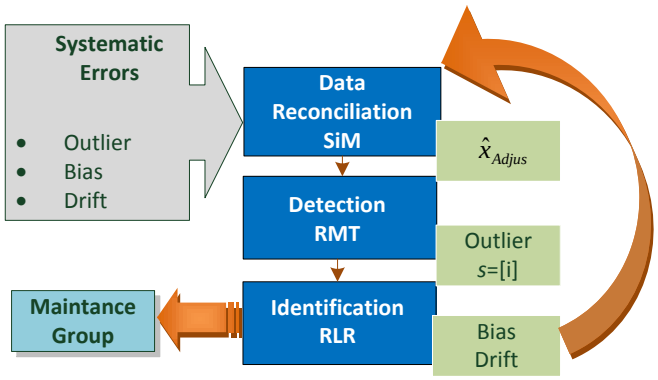

y_j
Y = [y_{j−N+1} ... y_j]
s=[ ]
Y* = f(Y, y_i^*)
n
y
DR
x̂_J^SIM
L_i=0
L_i=1
L_i=2
L_i=3
RMT
t̂_{i,j}^R = |a_{i,j}^R| / √Q̂_{ii}^R
t̂_{i,j-1}^R > t_{c1}
t̂_{i,j}^R > t_{c2}
Out=out+1
Oe_{i,j-1}
L_i=1
RMT
t̂_{i,j}^R
t̂_{i,j}^R > t_{c3}
Out=out+2
Oe_{i,j-2}
Oe_{i,j-1}
L_i=2
RMT
t̂_{i,j}^R
t̂_{i,j}^R > t_{c4}
Out=out+3
Oe_{i,j-3}
Oe_{i,j-2}
Oe_{i,j-1}
L_i=3
s=[i]
y_i^* = x̂_{J−4}^SIM
m=m+1
m=N/2
RLR
T_{β1,i} = β̂_{1,i}^R / [var(β̂_{1,i}^R)]^{1/2}
T_{β1} > T_{β1,c}
Bias=Bias+1
L_i=4
Drift=Drift+1
Li=5
J=j+1
Maintenance
P_i=0
s=[ ]
L_i=0
L_i=4
m=m+1
m=N
B_{i,2}
L_i=0
y_i^*=b_i


Y_{ob}^{N-2}
Y_{ob}^{N-3}
1 outlier
Y_{ob}^{N-4}
2 outliers
Y_{ob}^{N-5}
3 outliers
SEPT
N
I


RMT
RLR
95%
Isolated Errors
Outliers
Detected: %OC
Not Detected
Systematic Errors Simulated
Outliers
Bias
Drift
50%
Biases
Detected %BD
Not Detected %BN
Correctly Classified: %BCC
Wrong Classified: %BWC
5%
Systematic continuos errors
50%
Drift
Detected %DD
Not Detected %DN
Correctly Classified: %DCC
Wrong Classified: %DWC


10
4
12
9
25
11
17
16
22
4
3
1
3
5
15
8
24
1
5
2
18
23
6
13
14
2
6
10
27
9
7
19
7
21
20
26
11
8
28


B2
C2
A3
A4
A5
B1
C1
A1
A2
A8
D2
B3
A6
A7
D1


Systematic Errors
Outlier
Bias
Drift
Data Reconciliation SiM
$\hat{x}_{Adjus}$
Detection RMT
Outlier s=[i]
Identification RLR
Bias Drift
Maintance Group

CLASSIFICATION OF SYSTEMATIC MEASUREMENT ERRORS WITHIN THE FRAMEWORK OF ROBUST DATA RECONCILIATION

Claudia E. Llanos[1], Mabel C. Sánchez[1], Ricardo A. Maronna[2]*

[1]Planta Piloto de Ingeniería Química (Universidad Nacional del Sur - CONICET), Camino La Carrindanga km 7, 8000, Bahía Blanca, Argentina

[2]Departamento. de Matemática - Facultad de Ciencias Exactas –Universidad Nacional de La Plata, La Plata 1900, Argentina

KEYWORDS. Systematic Measurement Errors, Data Reconciliation, Robust Statistics

ABSTRACT. A Robust Data Reconciliation strategy provides unbiased variable estimates in the presence of a moderate quantity of atypical measurements. But estimates get worse if systematic measurement errors that persist in time (e.g., biases, drifts) are undetected and the break down

point of the robust strategy is surpassed. The detection and classification of those errors allow taking corrective actions on the inputs of the Robust Data Reconciliation that preserve the instrumentation system redundancy while the faulty sensor is repaired. In this work, a new methodology for variable estimation and systematic error classification, which is based on the concepts of Robust Statistics, is presented. It has been devised to be part of the real-time optimization loop of an industrial plant, therefore it runs for process operating under steady state conditions. The Robust Measurement Test is proposed in this article and used to detect the presence of sporadic and continuous systematic errors. Also the Robust Linear Regression of the data contained in a moving window is applied to classify the continuous errors as biases or drifts. Results highlight the performance of the proposed methodology to detect and classify outliers, biases and drifts for linear and nonlinear benchmarks.

1. Introduction.

Measurements are subject to errors that generate inconsistencies with the plant conservation equations. To reduce the detrimental influence of random measurement errors on variable estimates, Classical Data Reconciliation (DR) methodologies have been proposed. They provide variable estimates that minimize the Weighted Least Square (WLS) estimator of the observation adjustments and satisfy the set of constraints that represent the process operation. Because inferences are based on the idea that random errors are drawn from a normal distribution, the presence of systematic errors (e.g., outliers, biases, drifts, etc.) leads to inaccurate estimates. Therefore diverse strategies have been developed to deal with that problem. They detect, identify and estimate the magnitudes of the systematic errors. A review of those methodologies can be found elsewhere.[1,2]

Contrary to classical statistical approaches, Robust Statistics aims at providing good estimates when data are extracted from a neighborhood of a probability distribution model. Therefore, if a robust estimator is used as objective function of the DR problem instead of the WLS one, unbiased estimates can be calculated even in the presence of a moderate quantity of atypical observations.[1,3]

Atypical measurements may be sporadic. Their errors, called outliers, are due to poor electrical connections and electromagnetic interferences, or instantaneous loads associated with slamming waves or slippage of a mechanical fixing. In contrast, other atypical observations persist in time. For instance, biases and drifts are time-continuous measurement errors caused by sensor calibration failures. While the bias magnitude is constant in time, the drift one varies in accordance with a certain function of the time.

Many researchers have proved the advantages of using robust M-estimators when measurements are only contaminated with outliers. In this sense, the Contaminated Normal,[4] Cauchy, Logistic, Lorentzian, Fair[5] and Hampel[6] functions were applied to five benchmarks and two industrial plants operating at steady state.[7] For comparison purposes, the objective functions were tuned to obtain the same relative efficiency for the ideal distribution. Promising results were attained using the Cauchy and Hampel estimators. Furthermore two procedures, which combined the strengths of redescending and monotone M-estimators,[8] were presented. They were called The Simple Method (SiM) and the Sophisticated Method (SoM), and their computational costs were low in comparison with the requirements of the Hampel estimator. Also, the Quasi Weighted Least Square (QWLS) function[9] was presented and successfully applied to interesting industrial examples. Recently, a comprehensive study about the performance of robust DR methodologies showed the efficiency of SiM.[10]

Furthermore the improvement in the accuracy of estimates obtained using the Correntropy (CO), QWLS, Fair and Hampel functions as M-estimators was analyzed[11]. In this study measurements were corrupted with outliers or biases. Also, the advantages of using the Hampel function plus an advanced step moving window for the resolution of dynamic DR problems was recently studied.[12] Measurements were only contaminated with biases or drifts.

All the aforementioned works were related to the robust estimation of measurements, but they did not consider the detection and classification of systematic errors that persist in time (SEPT), e.g., biases and drifts. This is an important issue because the early detection of SEPT avoids that the variable estimates get worse when the break down point of the robust strategy is surpassed. This is related to the maximum amount of atypical observations that the data may contain to still give some information about the distribution of the typical measurements.[3] Furthermore, the categorization of a SEPT as a bias and the calculation of its magnitude allow estimating the bias-free observation. Also the results of the continuous error classification are helpful for the instrumentation maintenance group.

Up to the present time, only two works have dealt with the robust classification of SEPT, but comprehensive performance analyses were not provided. Regarding the first one,[13] the estimation problem was formulated using the Welsch M-estimator and solved applying the Particle Swarm Optimization algorithm. The strategy coped with the identification of outliers and biases. A cut-off value of the standardized observation adjustment[7] was used to determine the presence of outliers. Biases were identified as a sequence of outliers of the same sign. The number of time intervals of that sequence was fixed using the 95% confidence level of the binomial distribution.

The methodology proposed in the second work[14] coped with the identification of outliers, biases and drifts and used the CO estimator as objective function. It was proposed to calculate a statistic for each measurement as function of its standardized adjustment and weight, and declare the presence of a non-random error if that statistic was greater than a critical value. A distance-time step criterion was introduced to distinguish between outliers and SEPT. A threshold of the sample variance of the observation adjustments was used to discriminate between biases and drifts, but no advice about how to fix that threshold was provided.

In this work, a new methodology for variable estimation and systematic error classification is presented. It has been devised to take part of the real-time optimization loop of an industrial plant, therefore it runs for process operating under steady state conditions. The Robust Measurement Test is defined in this article and used to detect the presence of outliers and SEPT. Also the Robust Linear Regression of the data contained in a moving window is applied to classify the SEPT as biases or drifts. A comprehensive analysis of the strategy performance is provided in terms of the Percentage of Total Detection of SEPT, the Percentage of False Alarms of SEPT, the Percentage of Total Detection of Systematic Errors, the Mean Square Error, and the Percentages of Detection and Correct Classification of the aforementioned systematic errors for linear and non-linear benchmarks.

The paper is structured as follows. In Section 2, the estimation problem is formulated. The new strategy is presented in Section 3. Next, the performance results of the proposed methodology are presented and discussed. A Conclusion section closes the article.

2. Problem Formulation

Measurements are subject to errors that generate inconsistencies with the plant conservation equations. The unavoidable presence of random errors from unknown sources can be represented by the following measurement model,

$$y_{ij} = x_i + e_{ij}, \quad (1)$$

where x_i is the unknown true value of the i-th variable, y_{ij} stands for its observation at the j-th time period, and e_{ij} symbolizes the unobservable independent random error of that observation. It is commonly assumed that it follows the normal distribution function and has zero mean and known standard deviation σ_i. In addition to random errors, measurements may be contaminated with systematic errors. In this work, the possible presence of outliers, biases and drifts is considered. If they occur once at a time, the previous measurement model can be reformulated by adding the corresponding systematic error.

Frequently atypical observations are sporadic, but there are others which persist in time until the faulty sensor is repaired. The first ones are contaminated with systematic errors called outliers, whose magnitudes are many times bigger than σ_i. Their presence affects the tails of the error density function, which tend to zero more slowly than the corresponding ones to the normal distribution. In this case the measurement model can be reformulated as follows

$$y_{ij} = x_i + e_{ij} + O_{ij}, \quad (2)$$

where O_{ij} indicates the magnitude and sign of the outlier at the *j*-th time interval, which is expressed as $O_{ij} = K_{ij}^{o}\ \sigma_i$, where K_{ij}^{o} is a constant.

Regarding the atypical observations that persist in time, their errors may follow different deterministic behaviors[2]. Some abnormal measurements are contaminated with systematic errors of constant magnitude called biases. In this case, the observation accuracy is deteriorated but its precision remains equal to σ_i. The measurement model is represented by the following equation

$$y_{ij} = x_i + e_{ij} + B_{ij}, \tag{3}$$

where B_{ij} is the bias magnitude, which is usually formulated as $B_{ij} = K_i^{b}\ \sigma_i$ and K_i^{b} is a constant.

Other atypical observations that persist in time are corrupted by systematic errors called drifts. Equation 4 shows the model of this type of observations, where m_{drift} and $f(t)$ are a constant and a function of time, respectively. Recently, linear[14] and quadratic[12] functions have been used to characterize the time dependence of that measurement error. It has non zero mean and its variance becomes greater than σ_i when time increases.

$$y_{ij} = x_i + e_{ij} + m_{drift}.f(t), \tag{4}$$

Diverse model-based methodologies have been developed to reduce the detrimental effect of systematic errors on variable estimates. Among them, performance studies have demonstrated the efficiency of robust DR when a fixed amount of outliers are contained in a moving data window.[7] That methodology comprises the statement and resolution of the following optimization problem[1]

$$[\hat{x}_j^R, \hat{u}_j^R] = \underset{x_j}{Min} \sum_{p=j-N+1}^{j} \sum_{i=1}^{I} \rho\left(a_{ip}\right) \quad s.t. \quad f\left(x,u\right)=0 \quad , \tag{5}$$

where

$$a_{ip} = \frac{y_{ip} - x_{ij}}{\sigma_i} \, , \tag{6}$$

and vectors $[\hat{x}_j^R, \hat{u}_j^R]$ represent the state of the process at the *j*–th time interval. This is defined as the solution of the optimization problem that minimizes the loss function ρ of the standardized observation adjustments a_{ip} for a data window of length N and satisfies the process model $f(x,u)=0$, where x and u are the vectors of measured and unmeasured variables of dimension I and U respectively.

Furthermore, the derivative of the loss function ρ with respect to the standardized observation adjustment a is called Influence Function (IF), and it is represented by ψ, i.e., $\psi = \rho'$. The weight function W is related to the IF as follows

$$W(a) = \begin{cases} \psi(a)/a & \text{if } a \neq 0 \\ \psi'(0) & \text{if } a = 0 \end{cases} \tag{7}$$

In general the following types of M-estimators have been used as loss functions for Problem 5:

a) Monotone: ρ is a convex function, therefore it is unbounded; ψ is an increasing function (Huber, Fair, QWLS)

b) Redescending with unbounded ρ: ψ tends to zero at infinity (Cauchy, Logistic, Lorentzian, Welsch, CO)

c) Redescending with bounded ρ; ψ=0 for values greater than a threshold (Hampel and Biweight, which is denoted as BW)

Monotone estimators have the advantage that the solution of Eq. 5 has a unique local minimum. Therefore the values used to start the iterative solution procedure may influence the number of iterations but not the final outcome. On the other side, they are sensitive to very large outliers, and therefore may have a low efficiency for heavy-tailed error distributions[3], i.e., distributions whose density tails tend to zero more slowly than the normal density tails. Redescending estimators (b and c) can be made very efficient for heavy-tailed data, but Problem 5 may have several local minima, which requires a good starting point to ensure attaining the "good" solution. Estimators with bounded ρ (c) completely reject large outliers.

The breakdown point of an estimate should be taken into consideration when robust methodologies are applied. Roughly speaking, "the breakdown point of an estimate $\hat{\theta}$ of the parameter θ is the largest amount of contamination (proportion of atypical points) that the data may contain such that $\hat{\theta}$ still gives some information about θ, i.e., about the distribution of the typical points". [3] Let us analyze what happens if a robust M-estimator is used for DR and SEPT are present. A well-known benchmark [15] is used with this purpose, whose flowchart is represented in Figure 1. It comprises 4 units and 7 streams, and only mass balance equations are considered as constraints of Problem 5. Therefore, the number of variables of the optimization problem is *I*=7.

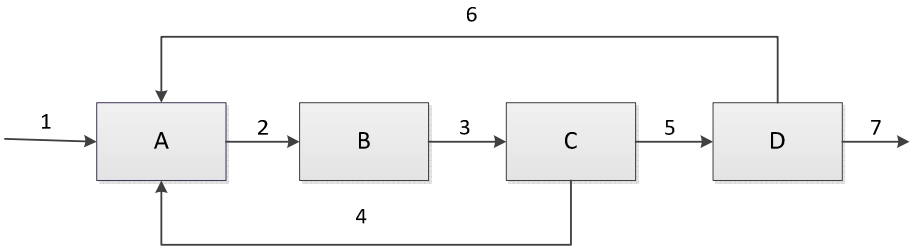


Figure 1. Case Study Flowsheet

Different types of M-estimators and measurement models are used. The efficiencies of the estimators are fixed at 95.5% for the ideal distribution by properly tuning their parameters. [7,8] The following measurement models are examined:

1) Model without systematic errors. It is assumed that random errors follow the standard normal distribution (Eq. 1)

2) Model with outliers. With probability $\varepsilon = 0.05$, an outlier of magnitude $O_{ij} = K_{ij}^{o}\,\sigma_i$ ($K_{ij}^{o} = 10$) is added to the random error (Eq. 2). The total number of generated outliers is 3457.

3) Model with SEPT. They are added to the random errors during 100 time intervals. The total number of simulated atypical observations is 3600, which is a similar amount to the one used for the previous model. The probability of occurrence of biases and drifts is the same. The magnitude of the biases is fixed at $K_i^b = 6$ (Eq. 3). A linear time dependent behavior is considered to simulate the drifts and m_{drift}=1(Eq. 4).

Furthermore, the standard deviations of the measurements are 2.5% of the true mass flowrate values, ten thousand simulation runs of the estimation problem are performed given an initial

data window, and the length of the data horizon is N=20. The Mean Square Error (MSE) is estimated as follows

$$MSE=\frac{1}{I\ J}\sum_{j=1}^{J}\sum_{i=1}^{I}\left(\frac{\hat{x}_{ij}^{R}-x_i}{\sigma_i}\right)^2, \qquad (8)$$

where J is the number of simulation runs. The MSE values obtained using different M-estimators and measurement errors are displayed in Table 1. It can be seen that all M-estimators overcome the detrimental effect of outliers on variable estimates; but the presence of SEPT degrades the estimation accuracy. Therefore, the application of a new strategy devoted to detect systematic errors, classify them, and take appropriate corrective actions when SEPT are present is worthwhile to enhance the accuracy of the estimates. That strategy is presented in the next section.

Table 1. MSE for different M-estimators and measurement models

	MSE		
M-estimators	Random	Outliers	SEPT
Huber	0,03114	0,03596	0,51394
CO	0,03111	0,03596	0,51396
Welsch	0,03118	0,03596	0,51406
BW	0,03118	0,03592	0,51392
Hampel	0,03118	0,03595	0,51392

3. Data Reconciliation and Systematic Error Classification

The proposed strategy interconnects three main procedures: robust DR, RMT, and RLR. In this section, they are described at first. The rationale of their interconnections is explained then.

3.1 Robust Data Reconciliation

In the last decade different strategies have appeared to solve the robust DR problem. A recent comparison among their performances[10] showed that SIM[8] can be applied for measurement adjustment in real time optimization loops, because it provides a good balance between the estimates accuracy and the computational load. Therefore, SIM is one of the main procedures of the proposed methodology.

The SIM is made up of two sequential steps which take advantage of the main features of redescending and monotone M-estimators:

Step 1: At the *j*-th time interval, a robust median of the *i*-th variable, $\tilde{y}_{ij}^{R}$ (*i*=1: *I*), is estimated using the measurements contained in a data window of length *N* $\left\{y_{ip}, p = j-N+1,...,j\right\}$. In this way, the redundancy supplied by the repeated observations allows to down weight the effect of atypical observations. The BW function is used to formulate the estimation problem

$$\hat{y}_{ij}^{R} = \underset{y_{ij}}{Min} \sum_{p=j-N+1}^{j} \rho_{BW}\left(\frac{y_{ip} - y_{ij}}{\sigma_i}\right) \tag{9}$$

where

$$\rho_{BW} = \begin{cases} 1-\left[1-(a/c_{BW})^2\right]^3 & if \quad |a| \le c_{BW} \\ 1 & if \quad |a| > c_{BW} \end{cases} \tag{10}$$

The iterative procedure is initialized using the solution of the previous DR problem.

Step 2: The state of the system at the j-th time interval is obtained by solving the following optimization problem using the solution of Step 1 as initial point

$$[\hat{x}_j^{SIM} \quad \hat{u}_j^{SIM}] = \underset{x_j}{Min} \quad \sum_{i=1}^{I} \rho_H \left(\frac{\hat{y}_{ij}^R - x_{ij}}{\sigma_i} \right)$$
$$s.t. \qquad (11)$$
$$f\left(x_j, u_j\right) = 0$$

where ρ_{H} corresponds to the Huber family

$$\rho_H = \begin{cases} a^2 & if \quad |a| \le c_H \\ 2c_H|a| - {c_H}^2 & if \quad |a| > c_H \end{cases} \qquad (12)$$

Because Problem 11 is defined in terms of a monotone estimator, its solution $[\hat{x}_j^{SIM}, \hat{u}_j^{SIM}]$ is unique. Furthermore the computation time of the iterative procedure is reduced because a good starting point is used.

Even though robust DR strategies are able to manage a moderate quantity of outliers, the presence of SEPT deteriorates the accuracy of estimates, as it was shown in the previous section. Therefore, it is worthwhile to detect and classify SEPT, and perform adequate corrective actions that enhance estimates quality.

3.2 Robust Measurement Test

The testing of statistical hypothesis has been widely used for the detection of atypical observations in DR problems. [1,2] In particular, the Measurement Test (MT)[16] was used in the past to point out the observations which may be corrupted by systematic errors. To briefly review its

application, let us consider that the mass flowrates of a process are all measured, and their observation adjustments are obtained using the data contained in a data window of length *N*. The vector of adjusted flowrates at the time interval *j*, $\hat{x}_j^{WLS}$, is the solution of Problem 13, which uses the WLS estimator as loss function:

$$\begin{aligned} \hat{x}_j^{WLS} = \underset{x_j}{Min} \quad & \sum_{p=j-N+1}^{j} \sum_{i=1}^{I} \left(\frac{y_{ip} - x_{ij}}{\sigma_i} \right)^2, \\ s.t. \quad & \\ & Gx = 0 \end{aligned} \tag{13}$$

where *G* represents the process incidence matrix. In consequence, the vector of measurement adjustments a_j^{WLS} is defined as

$$a_j^{WLS} = y_j - \hat{x}_j^{LS} = y_j - (\bar{y}_j - \Sigma G^T \left(G \Sigma G^T \right)^{-1} G \bar{y}_j) = y_j - (\bar{y}_j - Z \bar{y}_j) = y_j - (\Im - Z \bar{y}_j) \tag{14}$$

where $\bar{y}_j$ is the average of the measurement vectors $\left\{ y_p, p = j - N + 1, ..., j \right\}$, Σ is the known diagonal covariance matrix of the observations, and $\Im$ stands for the identity matrix. If only random errors are present $a_j^{WLS} \square N(0, Q)$ where

$$Q = \left[\frac{(N-1)\Im + Z}{N} \right] \Sigma \left[\frac{(N-1)\Im + Z}{N} \right]^T + \left[\frac{(N-1)}{N^2} \right] (\Im - Z) \Sigma (\Im - Z)^T \tag{15}$$

Using the previous results, the univariate statistic $t_{i,j}^{WLS}$ was defined for testing the *i*-th observation at the *j*-th time interval.[16]

$$t_{i,j}^{WLS} = \frac{a_{ij}^{WLS}}{\sqrt{Q_{ii}}} \tag{16}$$

It follows the standard normal distribution $N(0,1)$ if no atypical observations are present in the data window. The occurrence of a systematic error for the i-th observation is pointed out if the absolute value of the statistic is greater than a threshold.

Because the breakdown point of the WLS estimate is close to zero, even a single measurement with systematic error is enough to invalidate the basis of the methodology causing smearing.[3,7] To take better decisions regarding the occurrence of a systematic error, the RMT is proposed in this work.

The statistic of the RMT is defined as follows

$$t_{i,j}^{R} = \frac{a_{ij}^{R}}{\sqrt{Q_{ii}^{R}}} \tag{17}$$

where the vector of robust measurement adjustments is stated as

$$a_{j}^{R} = y_{j} - \hat{x}_{j}^{SIM} \tag{18}$$

and Q^R is its covariance matrix. If typical observations are present in the data window and the process operation can be represented by a linear set of equations, a_j^R follows approximately a $N(0,\ Q^R)$. This occurs because $\hat{x}_j^{SIM}$ is estimated using a linear transformation of the robust median, and the asymptotic normality of this M-estimate of location has been rigorously demonstrated. [3] The distribution of a_j^R is unknown if the process model is nonlinear. It will be explained later how to proceed in this case.

Since Q^R is unknown, a robust estimation of this matrix, $\hat{Q}_j^R$, is calculated at time j. First, the matrix A_j^R is formed containing the last a_p^R vectors (p=j-N+1,…, j),

$$A_j^R = [a_{j-N+1}^R, a_{j-N+2}^R, ...a_j^R] \tag{19}$$

Then, the normalized median absolute deviation about the median, $MADN(a_i^R)$, is estimated using the *i*-th row of A_j^R as follows

$$MAD\left(a_i^R\right) = Med\left[\left|A_j^R(i,:) - Med\left(A_j^R(i,:)\right)\right|\right] \tag{20}$$

$$MADN\left(a_i^R\right) = \frac{MAD\left(a_i^R\right)}{0.675} \tag{21}$$

and the square of $MADN(a_i^R)$ (i=1…I) is included into the scale estimate vector $\hat{\sigma}_a^2$. Next the matrix $\hat{Q}_j^R$ is evaluated

$$\hat{Q}_j^R = \hat{\sigma}_{\mathrm{a}}^2 \left\{ \frac{ave\left[\psi(A_j^R)/\hat{\sigma}_{\mathrm{a}}\right]^2}{\left(ave\left[\psi'(A_j^R/\hat{\sigma}_{\mathrm{a}})\right]\right)^2} \right\}^T \tag{22}$$

where ψ is the derivative of ρ_{BW} and ave represents the sample average.

The statistic of the RMT, $t_{i,j}^R$, is reformulated using $\hat{Q}_j^R$ as follows

$$\hat{t}_{i,j}^R = \frac{a_{ij}^R}{\sqrt{\hat{Q}_j^R\big|_{ii}}} \tag{23}$$

It follows the Student distribution with a number of degree of freedom df=N-1 if a_j^R is asymptotically normally distributed, i.e. $\hat{t}_{i,j}^R \sim t_{N-1}$.[3] In this case, the critical value of $t_{i,j}^R$ for a level of significance of the test α is $\left|t_{N-1,\alpha/2}\right|$. If the distribution of a_j^R is unknown, the following procedure is recommended to estimate the statistic critical value. At first, it is assumed that a^R follows an asymptotic normal distribution. This is a working hypothesis that allows calculating

$|t_{N-1,\alpha/2}|$. Then this assumption should be validated using a sample of a^R. If the working hypothesis is rejected, i.e. the experimental probability of $\{|t_i^R| > |t_{N-1,\alpha/2}|\}$ is different from α, the sample is used to calculate the critical value of the statistic. For example, kernel density estimation techniques can be applied with this purpose.

3.3 Robust Linear Regression

Regression analysis is a statistical tool for estimating the relationships between 2 or more variables. Let us consider fitting the following straight-line regression model

$$y = \beta_0 + \beta_1 x \tag{24}$$

to the data set $\{(x_m, y_m)$: m=1…M$\}$, where x_m and y_m are the predictor and response variable values, respectively. The regression model can be posed in compact form as follows

$$y = X\beta \tag{25}$$

where: $\beta = [\beta_0 \quad \beta_1]'$, and each row of matrix X is formed by the vector X_m=[1 x_m] (m:1…M). Robust regression methods provide estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ that aim at giving a good fit to the bulk of the data without being perturbed by a small proportion of atypical measurements. The vector of regression M-estimates $\hat{\beta} = [\hat{\beta}_0 \ \hat{\beta}_1]$ is defined as the solution of the following optimization problem:

$$Min \quad \sum_{m=1}^{M} \rho\left(\frac{r_m\left(\hat{\beta}\right)}{\hat{\sigma}_r}\right), \tag{26}$$

where

$$r_m = y_m - \left(\hat{\beta}_0 + \hat{\beta}_1 x_m\right), \tag{27}$$

and $\hat{\sigma}$ is its scale estimation, that may be calculated as:

$$\hat{\sigma} = \frac{1}{0.675} Med\left(\left|r_m\right| \quad \left|r_m \neq 0\right.\right) \tag{28}$$

The necessary and sufficient condition for solving Problem (26) is:

$$\sum_{m=1}^{M} \psi\left(\frac{\hat{r}_m}{\hat{\sigma}_r}\right) X_m = 0 \tag{29}$$

Using Eq. (7), ψ can be substituted by the weight function W as follows

$$\sum_{m=1}^{M} W\left(\frac{\hat{r}_m}{\hat{\sigma}_r}\right) \hat{r}_m \; X_m = 0 \tag{30}$$

From Eq. (30), the following expressions are obtained for the regression model parameters[3]:

$$\hat{\beta}_0 = \frac{\sum_{m=1}^{M}\left[W\left(\frac{\hat{r}_m}{\hat{\sigma}_r}\right)\left(y_m - \hat{\beta}_1 x_m\right)\right]}{\sum_{m=1}^{M}\left[W\left(\frac{\hat{r}_m}{\hat{\sigma}_r}\right)\right]} \tag{31}$$

$$\hat{\beta}_1 = \frac{\sum_{m=1}^{M}\left[W\left(\frac{\hat{r}_m}{\hat{\sigma}_r}\right) x_m\left(y_m - \hat{\beta}_0\right)\right]}{\sum_{m=1}^{M}\left[W\left(\frac{\hat{r}_m}{\hat{\sigma}_r}\right) x_m^2\right]} \tag{32}$$

The variance of $\hat{\beta}$ is calculated as:

$$\mathrm{var}\left(\hat{\beta}\right) = \hat{v}\left(X^T X\right)^{-1} \tag{33}$$

where:

$$\hat{v} = \hat{\sigma}^2 \frac{ave\left[\psi\left(\frac{r_m}{\hat{\sigma}_r}\right)^2\right]}{\left(ave\left[\psi'\left(\frac{r_m}{\hat{\sigma}_r}\right)\right]\right)^2} \frac{M}{M-2} \quad (34)$$

3.4 New Robust Data Reconciliation Methodology

For each time *j* the strategy comprises three stages. In the first one, the robust DR problem is solved. The second stage is related to the classification of the errors, and the third one updates the inputs for the next DR problem if it is necessary.

The experimental probability that four consecutive $\hat{t}_i^R$ surpass the critical statistical value for α=0.025 when measurements are drawn from the normal distribution is zero for any *N* when ten thousand observations are generated. This fact is used to distinguish between sporadic and continuous systematic errors. Therefore, if the RMT detects this event for the *i*-th variable, the sensor is considered faulty. Its measurements are replaced during the next time intervals by the values generated using the DR solution obtained for the last normal observation. This avoids the contamination of the data window with atypical observations while other corrective actions are executed.

When *N/2* unusual measurements of the *i*-th variable are available, the RLR is applied to classify the previously detected SEPT as a bias or a drift, and the problem is reported to the maintenance group. If the SEPT has been classified as a bias, its magnitude is estimated when *N* atypical measurements are collected. The following estimation problems use the bias-free measurement until the sensor is repaired.

Figure 2 is the flowchart of the proposed strategy. Let us define a vector s of faulty sensors that includes the instruments for which a SEPT have been detected. In the first stage, the new observation vector y_j is incorporated into the measurement matrix *Y(I,N)* by dropping the oldest vector of observations out of the matrix and appending y_j as the last column. If s=[ ], that matrix is used as input of the DR problem. If it is not the case, Y is modified as it will be explained later, and the measurement matrix Y^* results. Then the SIM is run to obtain $\hat{x}_j^{SIM}$.

As can be seen in Figure 2, the tasks involved in the second and third stages depend on the values of the indexes L_i and P_i. (i =1,…, I). The first one summarizes the results obtained during the previous time intervals regarding the classification of a systematic error for the i-th variable. The second one is a binary variable that indicates if the sensor has been repaired (P_i =1) or not (P_i =0). In Table 2, the initial variable values are displayed.

Table 2. Initial variable values

Variable	Value
s	[ ]
L_i	0
P_i	1

For L_i in the range [0 1 2], the statistic $t_{i,j}^R$ is evaluated after the reconciliation stage. If L_i =0, an outlier is detected for the penultimate measurement when $t_{i,j-1}^R$ is greater than the critical value but $t_{i,j}^R$ does not exceed this limit. In contrast, if both $t_{i,j-1}^R$ and $t_{i,j}^R$ go beyond the critical value, there is not enough information to classify the set of two atypical observations, and L_i is set equal

to 1. Because the presence of one or two atypical observations does not affect the results of the robust DR for the values of *N* commonly used, it can be run next without changing its inputs.

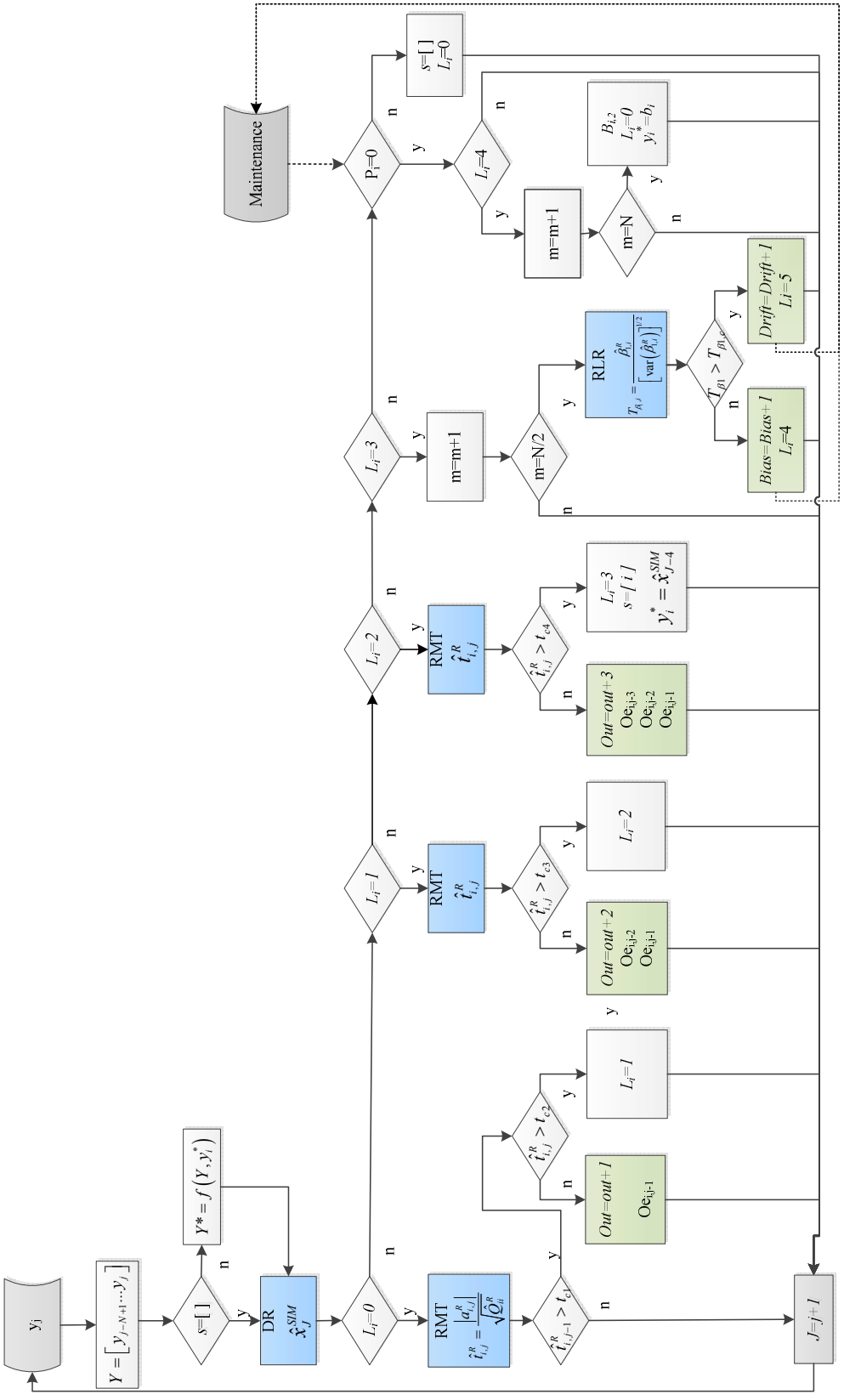


Figure 2 Flowchart of the proposed methodology

If L_i =2 and $t_{i,j}^R$ is not greater than the critical value, the presence of three consecutive outliers is identified, and L_i =0. In contrast, it is considered that the sequence of four unusual observations is part of a SEPT. In this case, the *i*-th variable is included in vector *s* and L_i is set equal to 3. Also *N* observations are generated using the adjusted value of the *i*-th variable obtained before the beginning of the SEPT ($\hat{x}_{i,j-4}^{SIM}$) plus random errors. These will form the *i*-th row of the matrix *Y** used for the next run of the DR procedure, denoted as y_i^*. Figure 3 illustrates a temporal sequence of observation vectors and shows how the presence of one outlier or consecutive ones is distinguished from the occurrence of a SEPT. The statistics whose values are smaller than the critical one are represented by a dot, otherwise they are symbolized using ⊗.

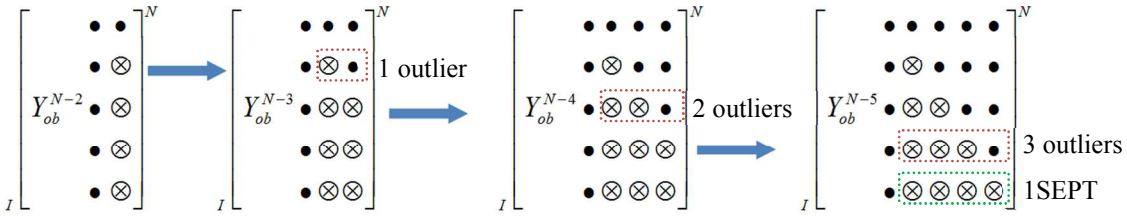


Figure 3: Detection of systematic errors

Once the presence of a SEPT has been declared for the *i*-th variable, their new measurements are not used for DR, and it makes no sense to calculate its statistic. These observations are saved until *M* consecutive atypical values are collected. At this moment, the RLR technique is run, and the estimate of the slope of the straight line is used to decide if the systematic error is a bias or a drift. The quantity of measurements used to estimate the straight-line regression model is controlled by *N* and the experimental break down point (*bdp*) of the estimate.[3] The following relation is considered:

$$M \leq bdp = \frac{N}{2} \tag{35}$$

Model parameters are estimated iteratively using the WLS solution as the initial value.

To classify the SEPT, the following statistical hypotheses are confronted:

H_0: $\hat{\beta}_1 = 0$

H_1: $\hat{\beta}_1 \neq 0$

and the statistic $T_{\beta 1}$, which considers the relation between $\hat{\beta}_1$ and its variance, is defined as

$$T_{\beta_1} = \frac{\hat{\beta}_1^R}{\left[\text{var}\left(\hat{\beta}_1^R\right)\right]^{1/2}} \quad \square \quad t_{df} \tag{36}$$

It follows the Student Distribution with *df=M-2*. If $T_{\beta 1} < T_{\beta 1,c}$, where $T_{\beta 1,c}$ is the critical value of the statistic for α=0.05, the SEPT is a bias, the parameter $\hat{\beta}_0$ represents its magnitude $B_{i,1}$, which is saved, and L_i is set equal to 4. On the contrary, the SEPT is a drift, and L_i is fixed at 5. When the classification of the SEPT finishes, an alert is sent to the instrumentation maintenance group changing the value of the binary variable P_i to 0. When the sensor is repaired, P_i will be equal to 1 as in the initial condition. No other tasks can be performed if a drift has been identified because measurements may follow diverse deterministic behaviors in time.

If the SEPT has been classified as a bias, new measurements are saved until a complete data window of observations corrupted by this error is available. At this moment, a new estimation of B_i, $B_{i,2}$, is calculated as the difference between the robust median of the measurements and the reconciled value of the variable, then L_i=0. The vector $b_i = B_{i,2}\ \text{ones}(N)$ is sent as input of the DR problem to correct the *i*-th row of Y_{j+1} as follows

$$Y^{*}_{j+1}(i,:) = Y_{j+1}(i,:) - b_i \tag{37}$$

until the sensor has been repaired. When this happens, P_i =1, then s=[ ], and L_i=0.

In Figure 2 the variables *out*, *bias* and *drift* are used to quantify the classification of the different systematic errors. Furthermore, the variable and time in which the algorithm performs the classification are informed. For example, $Oe_{i,j-1}$ refers to the presence of an outlier in the i-th variable at the (j-1)-th time interval.

4. Performance Analysis

A performance analysis of the developed methodology has been performed to quantify its capabilities for the detection and classification of systematic errors, and how these capabilities impact on the quality of the reconciled estimates.

Two benchmarks are studied: the Steam Metering Network (SMN) [17] and the Heat Exchanger Network (HEN)[18]. For each benchmark, four case studies are proposed that involve different values of the parameters N, B_i and m_{drift}, for a fixed magnitude of the outlier ($K_i^o = 10$) . Table 3 displays the parameter values and indicates if the proposed methodology has been applied or not.

Case I shows the application results of the technique when only random errors are simulated. In contrast, Case II presents its behavior when measurements are corrupted with outliers, biases and drifts. Case III is the worst condition for testing the methodology because the magnitudes of biases and drifts are smaller than those set for Case II. Also, it should be remarked that no articles in the literature have shown the behavior of techniques, which have the same purpose, when error magnitudes are so small. Regarding Case IV, it shows the estimation results when the proposed strategy is not applied and parameters are set as in Case II.

Table 3. Description of the case studies

Case	K_i^o	B_i	m_{drift}	Method
I	0	0	0	Yes
II	10	6	1	Yes
III	10	4.5	0.75	Yes
IV	10	6	1	No

Fifty thousand moving windows are considered for each case study. Most of the simulated observations are corrupted by random errors (Eq. 1). With a probability of occurrence p=0.02, the remaining measurements are randomly contaminated with systematic errors. This random procedure allows simulating the presence of consecutive outliers, as well as their simultaneous occurrence with biases or drifts. The 95% of the simulated systematic errors are outliers that last one time interval, and the rest of them are biases and drifts in the same proportion. When these appear, they persist during 100 time intervals, therefore 10% of the simulated observations are atypical values. Once a SEPT is detected, the corrective action lasts during 100 time intervals, after which it is considered that the sensor is available again. The procedure was executed using a Processor Intel ® Core (TM) i7 CPU 930 @ 2.80 GHz, 8GB RAM, using the Successive Quadratic Programming code of MatLab Release 7.12 (R2011a) to solve the nonlinear benchmark.

Global and individual performance indexes are proposed. The former ones are the Percentage of Total Detection of SEPT (%TD_{SEPT}), the Percentage of False Alarms of SEPT (% FA_{SEPT}), the Percentage of Total Detection of Systematic Errors (%TD) and the MSE. The first three measures are defined as follow:

$$\%\mathrm{TD}_{SEPT} = \frac{(\#biases + \#drifts)_{Simulated\ and\ Detected}}{(\#biases + \#drifts)_{Simulated}} x\,100 \quad (38)$$

$$\%\mathrm{FA}_{SEPT} = \frac{(\#SEPT)_{Detected} - (\#SEPT)_{Simulated\ and\ Correctly\ Detected}}{(\#SEPT)_{Detected}} x\,100 \quad (39)$$

$$\%\mathrm{TD} = \frac{(\#outliers + \#biases + \#drifts)_{Simulated\ and\ Detected}}{(\#outliers + \#biases + \#drifts)_{Simulated}} x\,100 \quad (40)$$

The $\%TD_{SEPT}$ indicates the percentage of simulated SEPT that are detected. Wrong detections of SEPT cause false alarms because the instrumentation maintenance group receives a message each time a SEPT is detected. These events are quantified using the $\%FA_{SEPT}$. The %TD represents the percentage of simulated systematic errors that are detected.

As it was shown in Section 2, the MSE takes into account the difference between the reconciled and true variable values. Even though a robust estimator is able to reduce the effect of a moderate amount of atypical observations, the presence of a SEPT biases the results of the robust DR problem if it is not detected, and the MSE grows. Therefore, it is useful to analyze the overall influence of the detection and classification procedure of SEPT on the reconciled variable values.

The individual performance indexes for the SEPT are evaluated in terms of their percentages of detection and correct and wrong classification. For outliers, just one index is computed. Figure 4 shows the relationship among those performance measures.

Simulated systematic errors are divided into Detected and Not Detected taking into account the results of the RMT. If isolated errors are detected, they are directly classified as outliers. In contrast, the RLR procedure is applied to categorize a detected SEPT as a bias or a drift.

Regarding outliers, the total set of detected outliers includes the simulated ones and those associated with the Type I Error selected for the RMT. The relation between the outliers which are correctly classified and the simulated ones is defined as the Percentage of Outlier Classification, %OC.

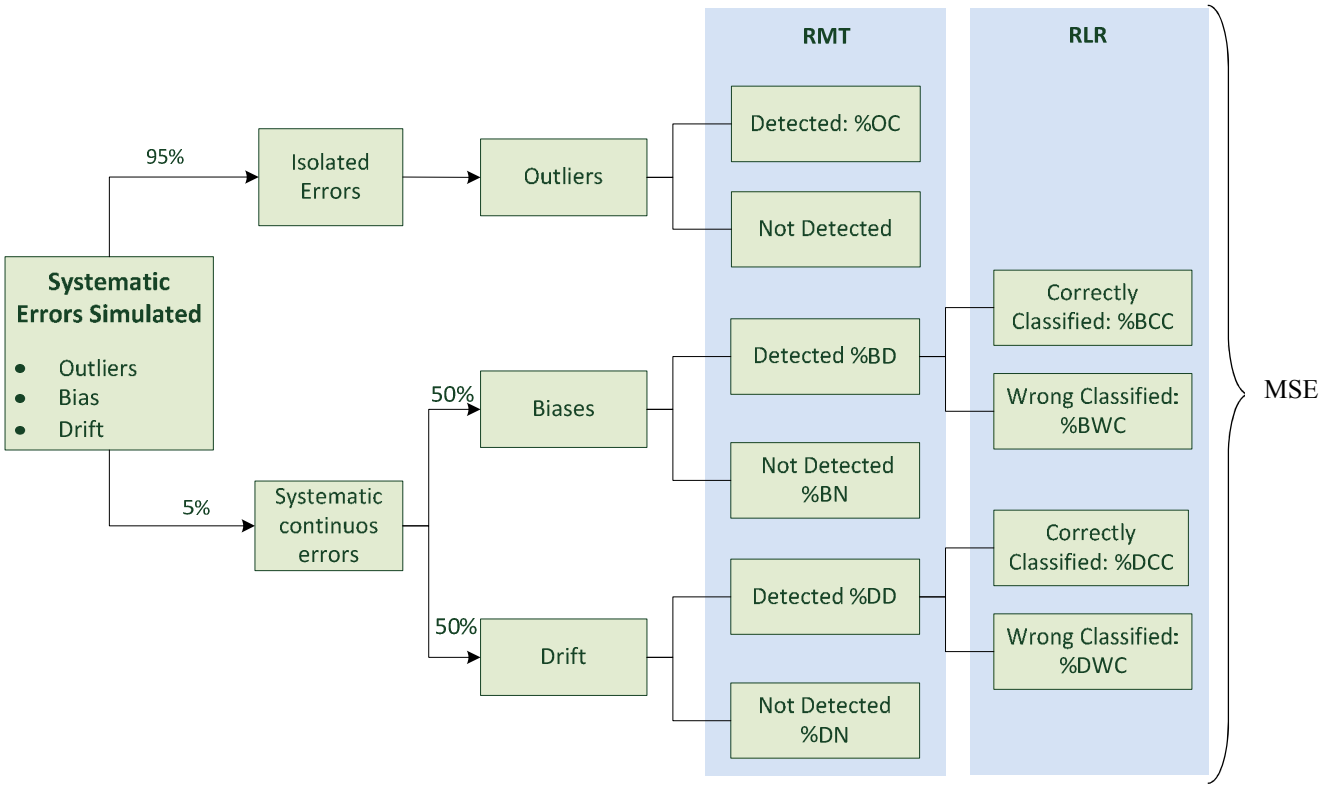


Figure 4. Relationship among individual performance indexes

With respect to biases, the detected ones take part of the Percentage of Detected Bias (%BD), which is divided into the Percentage of Correctly Classified Biases (%BCC) and the Percentage of Wrongly Classified Biases (%BWC). The same measures are defined for drifts. The

Percentage of Detected Drifts (%DD) is decomposed into the Percentage of Correctly Classified Drifts (%DCC) and the Percentage of Wrongly Classified Drifts (%DWC). That is:

$$\%OC=\frac{(\#outliers)_{Simulated\ and\ Correctly\ Classified}}{(\#outliers)_{Simulated}}x\,100 \tag{41}$$

$$\%BD=\frac{(\#biases)_{Simulated\ and\ Correctly\ Detected}}{(\#biases)_{Simulated}}x\,100 \tag{42}$$

$$\%BCC=\frac{(\#biases)_{Correctly\ Classified}}{(\#biases)_{Simulated}}x\,100 \tag{43}$$

$$\%BWC=\frac{(\#biases)_{Wrongly\ Classified}}{(\#biases)_{Simulated}}x\,100 \tag{44}$$

$$\%DD=\frac{(\#drifts)_{Simulated\ and\ Correctly\ Detected}}{(\#drifts)_{Simulated}}x\,100 \tag{45}$$

$$\%DCC=\frac{(\#drifts)_{Correctly\ Classified}}{(\#drifts)_{Simulated}}x\,100 \tag{46}$$

$$\%DWC=\frac{(\#drifts)_{Wrongly\ Classified}}{(\#drifts)_{Simulated}}x\,100 \tag{47}$$

To test the performance of a method for pointing out which observation is corrupted by a systematic error, it is a common practice[15] to generate an atypical observation vector and run the procedure of the strategy. The result of each simulation trial does not depend on the outcome of the previous trial. In contrast, the methodology proposed in this work performs different actions

that depend on the behavior in time of each "robust" measurement adjustment. Therefore it results necessary to develop an ad-hoc testing procedure for the new method.

Furthermore, information regarding the detection times (Dt) of SEPT is provided. In this sense, the median of the Dt for BCC (Dt_{BCC}), the number of BCC for which $Dt_{BCC}>N$ and the mean of Dt_{BCC} when $Dt_{BCC}>N$ are reported. The same values are presented for the BWC, the DCC and the DWC.

5. Analysis of Results

Next application results of the methodology to two benchmarks are reported and analyzed in detail.

5.1 Steam Metering Network (SMN)

The SMN represented in Figure 5 involves 28 streams that interconnect 11 units. The flow rates of all streams are measured. Random errors are generated considering that the standard deviation of the observations are 2.5% of their true values.

The performance measures are computed for Cases II and III, except the MSE that is evaluated for all the case studies. For Case II, the global and individual indexes are presented in Tables 4 and 5, respectively, and Table 6 displays the Dt for the SEPT. The same information is reported in Tables 7, 8 and 9 for Case III. In Table 10, the MSE values are displayed.

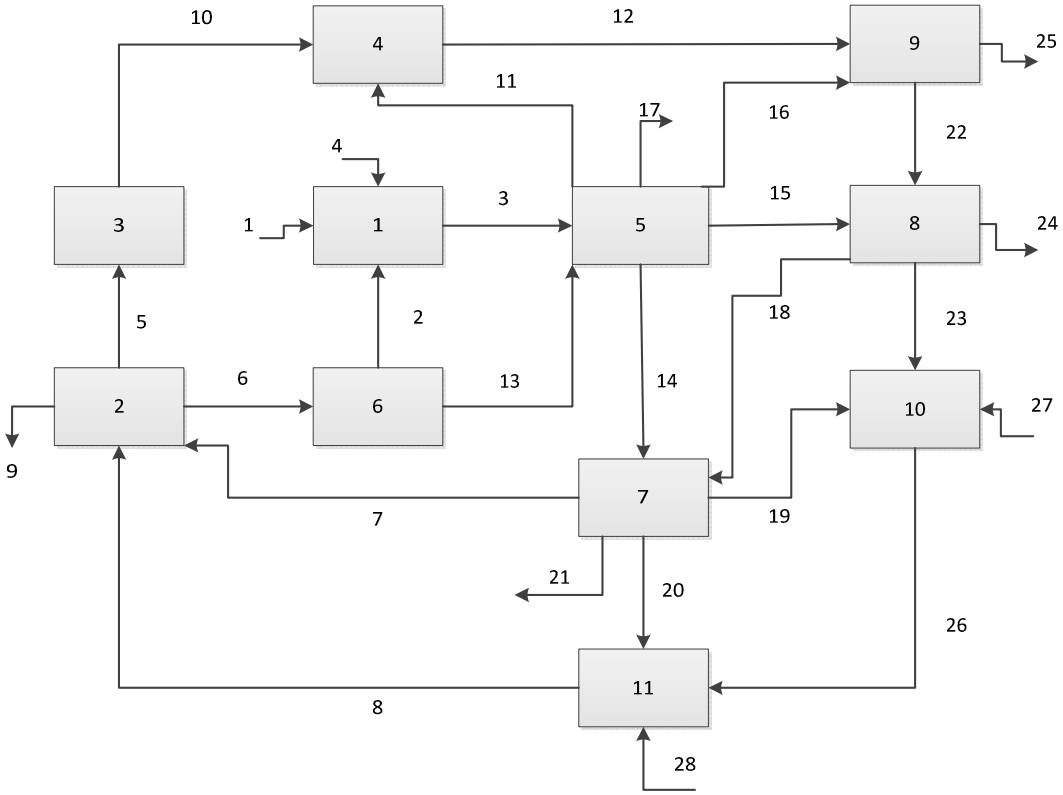


Figure 5. SMN Flowsheet

Table 4. Global Performance Indexes vs *N* - Case II (SMN)

N	% TD_{SEPT}	% FA_{SEPT}	% TD	MSE
24	92,55	11,10	91,91	126,36
30	96,58	6,68	95,59	22,11
40	95,44	15,87	93,06	45,93

Table 5. Individual Performance Indexes vs *N* - Case II (SMN)

N	% OC	%BD	%BCC	%BWC	%DD	%DCC	%DWC
24	91,87	93,28	86,87	6,42	91,80	90,71	1,08 (7)
30	95,53	97,02	88,96	8,06	96,13	95,67	0,46 (3)
40	92,93	95,52	90,75	4,78	95,36	94,12	1,24(8)

Table 6. Detection times vs N - Case II (SMN)

N	Dt_{BCC}			Dt_{BWC}			Dt_{DCC}			Dt_{DWC}		
24	4	11	61	4	1	81	7	2	74,5	57	7	53,6
30	4	5	50,2	4	0	---	6,5	1	37	25	1	40
40	4	5	55,8	4	2	81	6	2	53	36	3	51,7

The occurrence of 1316 SEPT, which last 100 time intervals, is simulated in different measurements. For Case II, the results obtained for $N<24$ are not included because their $\%TD_{SEPT}$ are lower than 90%. The *N* values are increased until the performance measure start to deteriorate. Next, these are reported and analyzed for three window lengths:

N =24: The highest MSE is obtained for this *N*, because the lowest $\%TD_{SEPT}$ and the highest Dt are achieved. If a SEPT is not detected, no corrective actions are performed, variable estimates get worse, then the variances of the measurement adjustments increase, and new atypical observations are not detected because their statistic values do not exceed the critical one. This has a negative effect on the %OC and the %TD, and they decrease. Furthermore the $\%FA_{SEPT}$ increases because if errors are not detected on time they may be noticed when they are finishing; in consequence a non-necessary corrective action starts.

N=30: Best percentages of total detection ($\%TD_{SEPT}$, %TD) and individual detection (%OC, %BD, %BCC, %DD and %DCC) are achieved. Regarding the Dt, SEPT are detected on time, thus the corrective actions are effective. Therefore the lowest MSE is obtained for Case II.

N=40: It is known that longer windows favor the increment of false alarms. This increase in the $\%FA_{SEPT}$ reduces the detection performance parameters, which are lower than the corresponding ones to *N*=30. The Dt for BCC, BWC and DWC are worse too. These differences cause that the MSE increase for *N* =40.

Table 7. Global Performance Measures vs *N* - Case III (SMN)

N	$\%TD_{SEPT}$	$\%FA_{SEPT}$	%TD	MSE
30	90,18	19,91	92,06	31,08
40	95,52	8,58	95,13	10,54
50	95,52	12,34	94,43	14,02

Table 8. Individual Performance Measures vs *N* - Case III (SMN)

N	% OC	%BD	%BCC	%BWC	%DD	%DCC	%DWC
30	92,16	87,46	80,15	7,31	93,03	91,80	1,24
40	95,11	94,03	89,40	4,63	97,06	96,28	0,77
50	94,37	94,18	89,70	4,48	96,90	95,98	0,93

Table 9. Detection times vs *N* - Case III (SMN)

N	Dt_{BCC}			Dt_{BWC}			Dt_{DCC}			Dt_{DWC}		
30	4	13	55,6	4	0	---	8	4	68,8	33,5	5	46,8
40	4	8	58,9	4	2	78,5	7	0	---	20	2	64,0
50	4	7	63,7	4	1	61,0	7	1	66,0	54	4	55,0

For Case III, the results obtained for N<30 are not included because their $\%TD_{SEPT}$ are lower than 90%. Next, the performance of the strategy is compared considering different values of N, denoted as N_1 and N_2, for $N \geq 30$:

N_1 =30 and N_2 =40. The percentages of total detection (% TD_{SEPT}, %TD) and individual detection (%OC, %BD, %DD) for N_1 are lower than the corresponding ones to N_2. The contrary happens for DWC. Also all the Dt are worse for N_1 in comparison to those obtained for N_2. In consequence, the MSE and $\%FA_{SEPT}$ are higher for N_1.

N_1 =30 and N_2 =50. In general results are similar to those discussed in the previous comparison, even though Dt_{BWC} is worse for N_2.

N_1 =40 and N_2 =50. The $\%TD_{SEPT}$ are equal for both window lengths, but the %TD is slightly lower for N_2 because the %OC diminishes due to the increment of false alarms. This happens because if a SEPT has been detected the execution of the RMT stops and outliers cannot be detected. Also the %DWC and Dt_{DWC} for N_2 increase with respect to the values obtained for N_1. This causes that the MSE and $\%FA_{SEPT}$ increase for N_2.

Table 10. MSE vs N (SMN)

N	I	II	III	IV
24	0,027	126,355	49.88	4943,537
30	0,022	22,113	31.08	4940,871
40	0,016	45,927	10.54	4283,548
50	0,013	19,834	14.02	3625,971

The MSE values obtained for Case I and Case IV can be considered as lower and upper bounds for this measure, respectively. Recall that Case I's observations are not contaminated with systematic errors while Case IV's measurements do. Because the proposed methodology has not been used for the last case, the presence of SEPT deteriorates the solution of the robust DR problem and the MSE increases. Furthermore, the MSE values for Case II show that variable estimates are significantly better than those achieved for Case IV. For Cases in which the methodology is used, acceptable performances measures are achieved for N=30.

5.2 Heat Exchanger Network

The operation of the HEN; which is illustrated in Figure 6. is represented by 17 mass and energy balance equations, which comprise 16 measured variables and 14 unmeasured ones. The standard deviation of flowrates and temperatures are 2% of their true values and 0.75K, respectively.

Global performances measures for Cases II and III are presented at Tables 11 and 14. The individual indexes are displayed in Tables 12 and 15, the Dt for different types of SEPT are included in Tables 13 and 16, and Table 17 presents the MSE for all the case studies.

Regarding Case II, satisfactory performance measures for both benchmarks are obtained for N=30. It should be noticed that the process model and the redundancy of the sensor network are different; this explains the differences on the MSE values attained for the benchmarks.

For the HEN, the best % TD_{SEPT} is achieved for N =24 and N =30. However the lowest MSE is obtained for N =30 because the % FA_{SEPT}, the individual performance parameters and the Dt are better for this N.

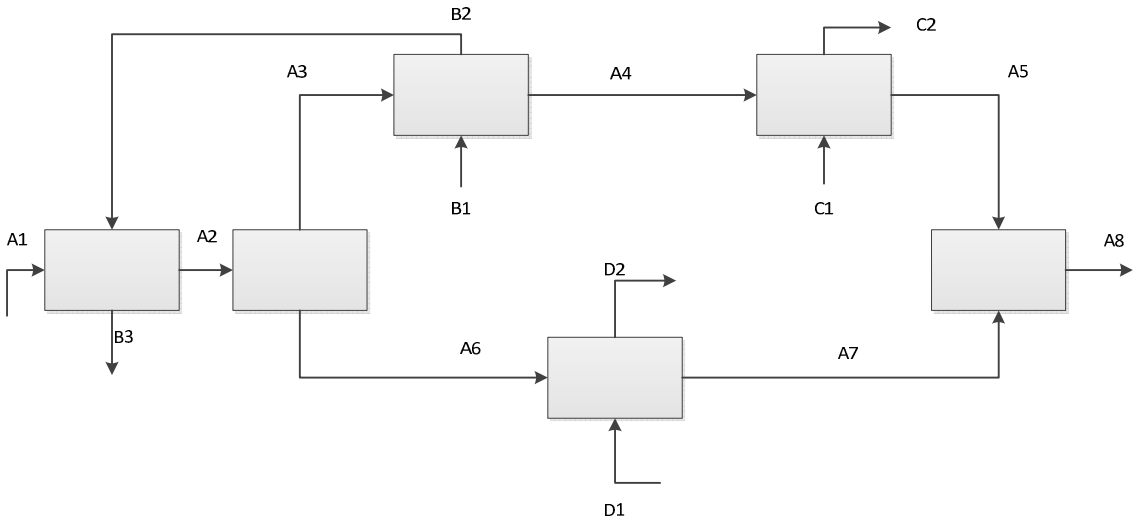


Figure 6. HEN Flowsheet

Table 11. Global Performance Indexes vs *N* - Case II (HEN)

N	$\%TD_{SEPT}$	$\%FA_{SEPT}$	%TD	MSE
24	96,43	11,70	96,19	6,20
30	96,43	7,63	96,59	2,12
40	95,19	13,91	94,77	14,26

Table 12. Individual Performance Indexes vs *N* - Case II (HEN)

N	%OC	%BD	%BCC	%BWC	%DD	%DCC	%DWC
24	96,179	96,364	90,909	5,455	96,501	96,210	0,292
30	96,599	95,325	88,571	6,753	97,668	97,668	0,000
40	94,746	93,506	87,273	6,234	97,085	96,501	0,583

Table 13. Detection times vs *N* - Case II (HEN)

N	Dt_{BCC}			Dt_{BWC}			Dt_{DCC}			Dt_{DWC}		
24	4	4	54.3	4	0	…	7	1	71.0	42	1	42.0
30	4	---	---	4	1	54.0	6	---	---	---	---	---
40	4	3	50.7	4	---	---	6	3	50.7	53.5	1	70.0

Regarding Case III, satisfactory performance measures for both benchmarks are obtained for *N*=40. With respect to the HEN, the best MSE is attained for *N*=40, and this measure decreases with the %TD. Some alarms of SEPT are declared just before the beginning of simulated SEPT. This increases the $\%FA_{SEPT}$ and diminishes the $\%TD_{SEPT}$, with respect to the values obtained for *N*=50, but the MSE is not significantly deteriorated because the errors are well classified and appropriate corrective actions are taken. Regarding *N*=50, the MSE increases because the number of simultaneous systematic errors contained in the window is larger, and this may reduce the accuracy of the estimates.

Table 14. Global Performance Indexes vs *N* - Case III (HEN)

N	$\%TD_{SEPT}$	$\%FA_{SEPT}$	%TD	MSE
30	89,56	20,68	92,34	60,37
40	93,41	11,57	95,77	7,47
50	94,09	8,67	95,39	13,19

Table 15. Individual Performance Indexes vs *N* - Case III (HEN)

N	% OC	%BD	%BCC	%BWC	%DD	%DCC	%DWC
30	92,481	86,494	80,519	5,974	93,003	90,962	2,041
40	95,897	90,130	82,857	7,273	97,085	95,918	1,166
50	95,456	91,169	87,013	4,156	97,376	97,085	0,292

Table 16. Detection times vs *N* - Case III (HEN)

N	Dt_{BCC}			Dt_{BWC}			Dt_{DCC}			Dt_{DWC}		
30	4	---	---	4	1	87	8	---	----	20	2	76,5
40	4	2	53,5	4	---	---	8	---	---	44,5	2	61,0
50	4	1	73,0	4	---	---	7	---	---	27	0	---

Table 17 shows the reduction on the MSE values obtained by applying the proposed methodology. As for the SMN benchmark, it is observed that longer windows are necessary to detect SEPT of small magnitude.

Table 17. MSE vs *N* (HEN)

N	I	II	III	IV
24	0,037	6,204	102,607	587,777
30	0,030	2,118	60,367	542,951
40	0,022	14,257	7,474	693,657
50	0,018	13,110	13,188	611,044

All the simulation studies indicate that if $N \leq 24$ the $\%TD_{SEPT}$ is low and the SEPT cannot be correctly detected and classified. For larger values of N, at least the 90% of the simulated SEPT are detected and 80% of the measurements contaminated with biases may be corrected. For $N=30$, an adequate relation between $\%TD_{SEPT}$, Dt and $\%FA_{SEPT}$ is achieved.

The proposed methodology has been applied to other case studies. For the sake of space, only some comments about the results obtained are included in this work next:

1. The performance of the strategy to deal with parallel streams is tested for process whose operation is represented by linear systems of equations. High percentages of detection and correct classification are obtained when SEPT are generated with the same probability for each observation. The same behavior is observed if SEPT are only simulated for the flowrates of the parallel streams.
2. The percentage of detection, correct classification, $\%TD_{SEPT}$ and MSE are evaluated when SEPT are simulated individually for each observation. Bias magnitudes and m_{drift} are changed in the range [0 -10] and [0 - 2], respectively, and f(t) is considered a linear function of time. As it is expected, results show that the performance measures for each measured variable increase with the Observation Redundancy Index. [19] This result confirms that DR methodologies provide better estimates when the system redundancy increases.
3. The proposed method can estimate the magnitude of the outlier as the difference between the measurement and the adjusted variable value obtained using the robust DR procedure. The estimation of the bias is also possible as it is previously explained. Furthermore, if it is assumed that the drift can be represented by a linear function of time, the slope can be estimated. But if this function is unknown, m_{drift} cannot be calculated.

4. In industrial scenarios process leaks can occur and may degrade the performance of the methodology.

Regarding the testing procedure of the method, it is considered that the occurrence of a moderate quantity of outliers does not significantly affect the accuracy of the estimates if robust DR procedures are applied. Thus, the false alarms due to the occurrence of outliers are not tested for Case I.

6. Conclusions

This work presents a new methodology for the detection and classification of systematic measurement errors. Its appropriate integration with a robust DR procedure significantly enhances the accuracy of variable estimates. The efficacy of the proposed algorithm is analyzed using two well-known benchmarks. In contrast to previous works, the type of systematic error and the time in which it occurs are randomly simulated. Furthermore, small error magnitudes are considered in this work.

The analysis of performance global measures indicates that the MSE diminishes with the increment of the %TD because appropriate corrective actions are taken to reduce the detrimental effect of systematic errors on variable estimates. Furthermore the occurrence of false alarms has an adverse effect on the detection of systematic errors and increases the MSE. Therefore, there is a tradeoff between the %TD and N because in general the $\%FA_{SEPT}$ increases with N.

The analysis of individual indexes demonstrates the importance of the right classification of SEPT. Their wrong categorization affects the MSE, especially if they are drifts. In these cases, erroneous corrective actions are taken and wrong information is provided to the DR procedure. This increments the false alarms, because once the corrective action ends, the RMT is applied

again to data windows containing measurements that present very different values. Some of them come from drifts wrong corrected and the other ones generally just contain random error. On the other hand, when biases are wrongly classified as drifts, the corrective action completely rejects the systematic error and good previous observations are considered as inputs of the robust DR; thus less error is introduced.

The individual percentages of detection and classification are useful to show the performance of the RMT and RLR procedures, respectively, while the MSE represents the accuracy of the attained variable estimates. However, all these performance measures are linked, because the MSE gets worse under the presence of undetected or wrongly classified SEPT. For this reason, not just the detection of SEPT is important, the correct classification is also necessary to reduce the MSE.

In this work, robust bias magnitudes are calculated. Therefore an 80% of the rectified observations of faulty sensors can be available until the sensor is repaired. As bias magnitudes are higher than those corresponding to the drifts, their estimates are more accurate and in consequence, the corrective action is more effective.

For the benchmarks and cases studies under analysis, high values of the percentages of detection and classification are obtained for N values in the range [30- 40]. Based on this results, it can be concluded that N=30 is an appropriated window length to apply quick corrective actions and achieve high performance measures.

The implementation issues of the method will be addressed in future works.

AUTHOR INFORMATION

Corresponding Author

*Mabel C. Sanchez. Tel.: +54 291 486 1700; E-mail address: msanchez@plapiqui.edu.ar

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors wish to thank the financial support of CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas), and UNS (Universidad Nacional del Sur, Bahía Blanca, Argentina).

REFERENCES

(1) Romagnoli, J.; Sánchez, M. *Data Processing and Reconciliation for Chemical Process Operations*; Academic Press: San Diego, 2000.

(2) Narasimhan, S.; Jordache, C. *Data Reconciliation and Gross Error Detection*; Gulf Publishing Company: Houston, 2000

(3) Maronna, R. A.; Martin, R.D.; Yohai, V. *Robust Statistics: Theory and Methods*; John Wiley and Sons Ltd.: Chichester, 2006.

(4) Tjoa, I.B.; Biegler, L.T. Simultaneous Strategies for Data Reconciliation and Gross Error Detection of Nonlinear Systems. *Comput. Chem. Eng.* **1991**, 15, 679-90.

(5) Albuquerque, J. S. ; Biegler, L.T. Data reconciliation and gross error detection for dynamic Systems. *AIChE J.* **1996**, *42*, 2841-2856.

(6) Arora, N.; Biegler L.T. Redescending estimators for data reconciliation and parameter estimation. *Comput. Chem. Eng*. **2001**, *25*, 1585-1599.

(7) Ozyurt, D.B.; Pike, R.W. Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes. *Comput. Chem. Eng*. **2004**, 28, 381-402.

(8) Sánchez, M.; Maronna, R. Simple Approaches for Robust Data Reconciliation. 2009 *AIChE Annual Meeting, Nashville, TN*. United States, **2009**, Code 79788.

(9) Zhang Z.; Zhijiang S.; Chen X.; Wang K.; Qian J. Quasi-weighted least squares estimator for data reconciliation. *Comput. Chem. Eng*. **2010,** 34*,* 154-162.

(10) Llanos C. E.; Sánchez, M. C.; Maronna, R. A. Robust estimator for Data Reconciliation. *Ind. Eng. Chem. Res.*, **2015**, 54, 5096–5105.

(11) Chen J, Peng Y, Munoz J. Correntropy Estimator for Data Reconciliation. *Chem. Eng. Sci*. **2013**, 104, 10019-10027.

(12) Nicholson, B.; López-Negrete, R.; Biegler, L. T. On-line State Estimation of Nonlinear Dynamic Systems with Gross Errors. *Comput. Chem. Eng*. **2014,** *70,* 149-159.

(13) Martinez Prata, D.; Schwaab, M.; Lima, E.L.; Pinto, J.C. Simultaneous Robust Data Reconciliation and Gross Error Detection through Particle Swarm Optimization for an Industrial Polypropylene Reactor. *Chem. Eng. Sci*. **2010,** *65,* 4943-4954.

(14) Zhang, Z.; Chen, J. Correntropy Based Data Reconciliation and Gross Error Detection and Identification for Nonlinear Dynamic Processes. *Comput. Chem. Eng*. **2015,** *75,* 120-134.

(15) Rosenberg, J; Mah, R S H; Iordache, C. ;Evaluation of Schemes for Detecting and Identifying Gross Errors in Process Data. *Ind. Eng. Chem. Res.*, **1987**, 26, 555–564.

(16) Mah, R. S. H. and Tamhane A. C. Detection of Gross Errors in Process Data. *AIChE J*., **1982**, 28, 5, 828-830.

(17) Serth, R.; Heenan, W. Gross error detection and data reconciliation in steam-metering systems. *AIChE J*. **1986**, 32, 733-41.

(18) Swartz, C.L.E.: Data reconciliation for generalization flowsheet applications. 197th *Natl. Meet*., *Am. Chem. Soc*., Dallas, TX (1989).

(19) Maronna, R.A.; Arcas, J. Data reconciliation and gross error detection based on regression. *Comput. Chem. Eng*. **2009**, 33, 65-71.