UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA INDUSTRIALE

CORSO DI LAUREA MAGISTRALE IN INGEGNERIA CHIMICA E DEI PROCESSI INUDSTRIALI

Robust data reconciliation and gross error detection in industrial heat exchanger networks

Relatore: Prof. Pierantonio Facco

Supervisori aziendali: Dr. Jose Loyola-Fuentes, Dr. Francesco Coletti

Laureando: TOMMASO ACERBI

ANNO ACCADEMICO 2019/2020

Abstract

Industrial measurements are invalidated by the presence of random errors (*i.e.*, background noise) and gross errors (*i.e.*, outliers, bias, etc…). Data reconciliation (Romagnoli and Sanchez, 2000) allows to reduce these effects by solving an optimization problem constrained to satisfy the mass and energy balances. However, the presence of gross errors deteriorates the reconciliation efficiency. In order to cope with this issue, the adoption of robust estimators is needed. Furthermore, it is possible detect and identify the gross errors through gross error detection methods (Romagnoli and Sanchez, 2000).

In this Thesis data reconciliation and gross error detection are exploited to provide reliable measurements for the estimation of fouling model parameters in crude-oil heat exchanger networks. In particular, six state-of-the-art robust estimators are applied to two simulated heat exchange systems (one static and one dynamic) and a real one with "cut-off points" and X84 (Ozyurt and Pike, 2004) for gross error detection. As a general result it is pointed out that Welsch estimator offers the best trade-off between data reconciliation and gross error detection performances; furthermore, the "cut-off points" proves to be the more suited gross error detection method when adopting robust M-estimators.

The work in this Thesis was done in collaboration with Hexxcell Ltd as a part of an industrial internship.

Riassunto

Le misurazioni industriali sono invalidate dalla presenza di errori casuali (*i.e.*, rumore di fondo) e da errori sistematici (*i.e.*, outlier, bias, ecc…). La riconciliazione dati (Romagnoli e Sanchez, 2000) è una tecnica che permette di ridurre questi effetti attraverso la risoluzione di un problema di ottimizzazione vincolata, soddisfacendo i bilanci di materia ed energia. Tuttavia, la presenza di errori sistematici riduce l'efficacia della riconciliazione dati. Per sopperire a questa problematica, l'adozione di stimatori robusti è necessaria per la riconciliazione. È possibile inoltre confermare ed indentificare la presenza di errori sistematici mediante metodologie specifiche.

La riconciliazione dati e l'individuazione degli errori sistematici sono sfruttate per fornire misurazioni affidabili nella stima dei parametri predittivi dello sporcamento nei sistemi di scambiatori di calore per il petrolio greggio. In particolare, sei stimatori robusti all'attuale stato dell'arte sono applicati a due sistemi simulati di scambio termico (uno statico ed uno dinamico) e ad uno reale. Allo stesso modo, le metodologie *"cut-off points"* e X84 (Ozyurt and Pike, 2004) sono applicate per individuare gli errori sistematici. In generale, si è concluso che lo stimatore di Welsch offre la miglior combinazione di prestazioni tra riconciliazione dati ed individuazione degli errori sistematici; inoltre, la metodologia dei *"cut-off points"* si è rivelata essere la più adatta per l'individuazione di errori sistematici, qualora si utilizzino gli stimatori robusti.

Il lavoro in questa Tesi è stato realizzato in collaborazione con Hexxcell Ltd come parte di un tirocinio aziendale.

Contents

Introduction

In industrial plants, the reliability of collected data has always been a challenge. Process monitoring, unit revamping or operational optimization are just ones of the most frequent situations where the process or systems reliability depends on the quality of the process measurements. Therefore, it is important to guarantee the validity and the accuracy of the collected and processed data. However, process data are inevitably corrupted by errors during collection and processing. This makes data deviate from their true values, leading to significant deviations from physical conservation laws (Tamhane and Mah, 1985).

Two types of measurement errors usually invalidate the measured data: random errors and gross errors (Miao et al., 2009). Random errors cannot be eliminated and are always present in the measurements. This is identified as random noise, and usually is of relatively limited magnitude, with the exception of few occasional spikes. This type of disturbance cannot be predicted certainly, therefore, the only possible way to characterize it is to use probability distributions. On the other hand, gross errors are due to non-random events. These errors occur less frequently, but their magnitude is typically larger. They can be due to process leaks, malfunctioning sensors and systematic bias. Since both random and gross errors lead to erroneous data, degrading the performances and safety of the whole process, it is necessary to minimise their impact on the measurements.

Data reconciliation (DR) is a well-known methodology that improvs the accuracy of process data by reducing the effect of random errors in measurements (Romagnoli and Sanchez, 2000). Typically, the reconciled estimates are expected to be more accurate than the measurements. However, the reconciled process data can suffer the presence of gross errors. For this reason, the effect of gross errors must be eliminated to guarantee the effectiveness of data reconciliation. Therefore, Gross Error Detection (GED) is a companion technique of data reconciliation that has been developed to identify and eliminate gross errors in process data (Romagnoli and Sanchez, 2000). DR and GED are applied together to improve the estimates accuracy of process data.

Over the past decades many techniques were developed to perform DR and GED. They can essentially follow two strategies:

- the strategies based on serial elimination (Ripps, 1965; Serth and Hennan, 1986; Rosenberg, Mah and Iordache, 1987), serial compensation (Narasimhan and Mah, 1987) and simultaneous/collective compensation (Sanchez, Romagnoli et al., 1999). All of them perform DR through a constrained WLS algorithm but then they detect and eliminate variables with gross errors using statistical test, such as Global Test, Measurement Test and Nodal Test (Tamhane and Mah, 1985).
- the strategies that mitigate the effect of gross errors by using cost functions that reduce the weight of variables with large errors. This are called robust estimators and do not detect, nor eliminate gross errors, but reduce their effect on the solution, thus avoiding the propagation of the error (Arora and Biegler, 2001).

Many authors pointed out the optimal performance of the robust estimators (Prata et al., 2010; Johnston and Kramer, 1995). For example, using Contaminated Normal estimators allow reconciled values to replace any outliers in the measured data, without requiring iterative detection and elimination procedures (Prata eta al., 2008). In particular, the performances of the M-estimators have been widely assessed (Ozyurt and Pike, 2004; Llanos et al., 2015), and a selection of six robust M-estimators was made in order to operate a comparison between the most promising estimators of the discussed over last decades: Simple Method, Sophisticated Method, Welsch estimator, Quasi Weighted Least Squares estimator, Correntropy estimator and Fair estimator. Along with these DR robust methods some effective GED techniques (such as, cut-off points and X84) were adopted (Ozyurt and Pike, 2004).

DR and GED techniques can have a specific application in thermal systems in order to improve the fouling modelling in crude oil pre-heat trains (Loyola-Fuentes and Smith, 2019). This issue is of great importance industrially because fouling causes additional fuel and production costs, difficult operations, CO_2 emissions and safety issues (Coletti and Macchietto, 2011). Fouling deposition can be predicted according to specific fouling rate models. These models are based on several parameters, which can be also estimated by measured process data. Therefore, having a confident data source is crucial, with reliable and free-of-error measurements.

In this perspective, the present Thesis deals with the application of robust DR and GED methods in the case of heat exchange. In particular, the performance of DR and GED methodologies proposed in Literature are tested and analysed in three case studies. The first two consist of simulated shell-and-tube heat exchangers, one treated as steady-state system and the second as a dynamic system. The third case study is related to an industrial heat-exchangers network (HEN) of a pre-heat train for crude-oil. From the simulated case studies, it is possible to evaluate the most robust methods for DR and GED.

In Chapter 1, the mathematical background behind DR, robust M-estimators and GED techniques is presented. Then the three case studies are illustrated in Chapter 2. Chapter 3 shows the discussion of the results in the preliminary simulated case studies. Chapter 4 deals with the application of DR and GED in the industrial system. Finally, in the concluding Chapter the final remarks and the future work are discussed.

CHAPTER 1

Mathematical background

In this Chapter the mathematical background regarding on robust data reconciliation (DR) and gross error detection (GED) techniques is presented. Starting from the general DR formulation and characteristics, the difference between steady-state DR and dynamic DR is highlighted. Then the robust M-estimators are explained. Finally, the formulations of the adopted GED techniques are presented and discussed.

1.1 Data reconciliation

Data reconciliation (DR) is a technique to improve the accuracy of process data by reducing the effect of random errors in measurements; by solving a constrained optimization problem, it adjusts of the process measurements to obtain more accurate estimates of flowrates, temperatures, compositions, etc…, that are consistent with material and energy balances.

Consider a process whose process variables are identified by I variables that can be measured with error. This set of system variables are stored in a vector:

$\mathbf{x}^{\mathrm{T}} = [x_1, x_2, x_i, \dots, x_I]$. Let us assume that a set of J (noisy) measurements $\mathbf{y}_{\mathrm{i}}$ are available as a function of the I variables in $\mathbf{x}$:

$$\mathbf{y}_i = f_i(x_i) + \boldsymbol{\varepsilon}_i \,, \tag{1.1}$$

where $\boldsymbol{\varepsilon}_i$ is the measurement error. These measurements are collected in vectors: $\mathbf{y}_i^{\mathrm{T}} = [y_{i1}, y_{i2}, y_{ij}, \dots, y_{iJ}]$ for $i = 1,2,3, \dots I$ and for $j = 1,2,3, \dots J$, where J is the total number of measurements taken during steady-state plant operation to estimate the system variable x_{i}: J = 1 if we are interested in the snapshot of the process; $J > 1$ if our concern is a smoothed average within a time window of interest.

If the difference of the measured values y_i and the system variable x_i is $\boldsymbol{\varepsilon}_i = 0$, the measurements are perfect and $\mathbf{y}_i = f_i(x_i)$. In general, in a state of statistical control and with only common cause variations occurring in the system (*i.e.*, without any assignable/special

cause), $\boldsymbol{\varepsilon}_i \neq 0$ and $\boldsymbol{\varepsilon}_i \sim \mathbf{e}_i = N\left(0, {\sigma_{\boldsymbol{\varepsilon}_i}}^2 \sim 1\right)$, namely, the measured values $\mathbf{y}_i$ and the system variable x_i do not coincide and the measurement errors are independent and normally distributed with zero mean. However, process data may contain other types of errors caused by non-random events (*e.g.*, assignable or special causes occurring, such as: instruments not adequately compensated, sensors malfunction, etc…). These are called gross errors and they can be occasional (*i.e.*, outliers) or systematic and time-persistent (*i.e.*, biases). Biases do not follow a particular distribution and their magnitude is usually higher with respect to the random errors (Miao et al., 2009). Under this situation Equation (1.1) can be rewritten as

$$\mathbf{y_i} = f_i(x_i) + \boldsymbol{\varepsilon}_i = f_i(x_i) + \mathbf{e}_i + \mathbf{o}_i \tag{1.2}$$

in the case outliers are present and

$$\mathbf{y}_i = f_i(x_i) + \boldsymbol{\varepsilon}_i = f_i(x_i) + \mathbf{e}_i + \mathbf{b}_i \tag{1.3}$$

in the case biases are present, where ${\mathbf{o}_i}^T$ and ${\mathbf{b}_i}^T$ are the $[1 \times J]$ vectors of variable i for outliers and bias magnitudes, respectively (Llanos et al., 2017).

In the case $\varepsilon_i \neq 0$ and f_i is differentiable at a point $x_{i,0}$, a vector $\mathbf{f}_i$ can be defined:

$$\mathbf{f}_i\left(x_{i,0}\right) = \left.\frac{df_i}{dx_i}\right|_{x_{i,0}} \quad , \tag{1.4}$$

which can be used for the linearized version of the measurements in Equation (1.1). For linear systems $\mathbf{f}_\mathrm{i}$ is constant and independent of x_i . Accordingly, a linearized system can be considered as a good approximation of Equation (1.1):

$$\mathbf{y}_i = \mathbf{f}_i x_i + \boldsymbol{\varepsilon}_i \quad , \tag{1.5}$$

where $\mathbf{f}_\mathrm{i}$ is the $[J \times 1]$ vector of the Jacobian of $\mathrm{f_i}$. To uniquely determine this model a minimum number I of independent variables exists. Defining as redundant a system in which the amount of available data exceeds the minimum amount necessary for a unique determination of the independent variables ($J > I$) (Romagnoli and Sanchez, 2000) that determine a chosen model, if the number of observations J is not sufficient for determining the I variables, the situation will be obviously deficient. Furthermore, note that not only the number of observations must be larger than the number of variables, but the observations must be also independent. This means that, when ($J > I$), redundancy exists and is defined as:

$$R = J - I \tag{1.6}$$

that is equal to the (statistical) degrees of freedom. The definition of redundancy through Equation (1.6) can be also expressed in terms of observability (Loyola-Fuentes and Smith, 2020) where:

- redundancy: a measured variable is defined as redundant if it is observable even when its measurement is not available.
- observability: an unmeasured variable is defined as observable if it can be estimated using the available process measurements and system equations (*i.e.*, material and energy balances).

From these two definitions it is clear that measured variables can be classified as redundant or non-redundant; the unmeasured variables, for which measurements are not available, can be only classified as observable or non-observable.

However, since the observations are obtained from measurements that are subject to probabilistic fluctuations, redundant data are usually inconsistent because each sufficient subset of measurement yields different results from other subsets. To obtain a unique solution, an additional criterion is needed: among all the solutions that are consistent with the measurement model, the estimates that are as close as possible to the measurements are considered to be the solution of the estimation problem. To this purpose a generalized least squares solution is considered (Romagnoli and Sanchez, 2000):

$$\hat{\mathbf{x}} = \min_{x_i} \sum_{i=1}^{I} (\mathbf{y}_i - \mathbf{f}_i x_i)\,(\mathbf{y}_i - \mathbf{f}_i x_i)^T. \tag{1.7}$$

The vector $\hat{\mathbf{x}}$ contains the final estimates which are the reconciled values of the measured variables and should be closer to the true values of the system variables.

System constraints are needed to satisfy material and energy balances, and depending on their formulation DR can be performed on a steady-state (c_{ss}) or dynamic system (c_{dyn}), as reported below,

$$c_{ss}(\hat{\mathbf{x}}) = \mathbf{0} \tag{1.8}$$

$$c_{dyn}\left(\frac{d\hat{\mathbf{x}}(t)}{dt}, \hat{\mathbf{x}}(t)\right) = \mathbf{0}\,. \tag{1.9}$$

1.1.1 Reduced data reconciliation problem

If some of the variables are not measured the DR problem can be reformulated into a reduced DR problem (Romagnoli and Sanchez, 2000). Consider that the unmeasured variables are

stored in the vector $\mathbf{u}^{\mathrm{T}} = [u_1, u, u_z, \ldots, u_Z]$. In this case, the dimensionality of the optimisation problem is reduced to reconcile only the redundant variables.

Let us assume a steady-state system with variables **k**. The problem constraints can be linearized according to the following equation:

$$\mathbf{Ak} = \mathbf{c}\,, \tag{1.10}$$

where **A** is the compatible matrix and ***c*** is the resulting vector from the equality constraints and it could equal to 0. A compatible matrix (Romagnoli and Sanchez, 2000) is a $[U \times I]$ which describes the structural topology of I streams and U units in terms of variables and equations. Consider for example the following simple system of Figure 1.1, whose mass balances around unit 1 and 2 are shown below,

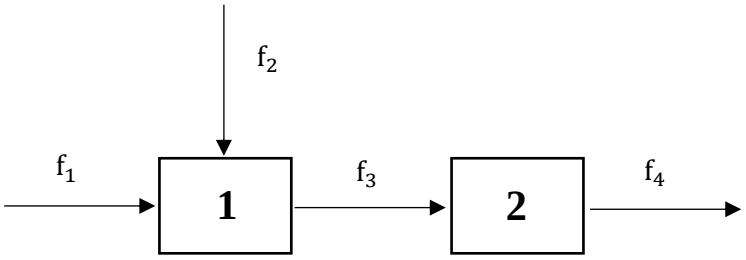


Figure 1.1. *Flowsheet diagram for a simple serial system.*

$$f_1 + f_2 - f_3 = 0$$
$$f_3 - f_4 = 0\,,$$

the correspondent compatible matrix **A** has the form:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

The previous **k** variables can be divided into **x** measured variables and **u** unmeasured variables. Therefore, Equation (1.10) changes to the following form:

$$\mathbf{A_x x} + \mathbf{A_u u} = \mathbf{c}\,, \tag{1.11}$$

where $\mathbf{A_x}$ and $\mathbf{A_u}$ are the compatible matrix respectively for the measured and unmeasured variables. The dimension of the overall problem is reduced by using a projection matrix **P** such that:

$$\mathbf{PA_u u} = \mathbf{0}\,,$$

obtaining the following reduced set of constraints,

$$\mathbf{PA_x x} = \mathbf{Pc}\,. \tag{1.12}$$

The projection matrix is estimated using the Q-R decomposition method (Romagnoli and Sanchez, 2000). Matrix $\mathbf{A_u}$ is decomposed using this method in matrices $\mathbf{Q_u}$ and $\mathbf{R_u}$. Depending on the case, a permutation matrix $\boldsymbol{\pi_u}$ can also be estimated with the purpose of reordering the columns of $\mathbf{A_u}$ if necessary. The factorisation results in:

$$\mathbf{A_u} = \mathbf{Q_u R_u \pi_u} = [\mathbf{Q_{u1}} \quad \mathbf{Q_{u2}}] \begin{bmatrix} \mathbf{R_{u1}} \\ 0 \end{bmatrix} \boldsymbol{\pi_u}\,, \tag{1.13}$$

and the matrix $\mathbf{Q_{u2}}^{\mathrm{T}}$ corresponds to the projection matrix **P** (Crowe et al., 1983). Using Equation (1.13) in combination with Equation (1.11), the solution for the unmeasured variables **u** is

$$\boldsymbol{\pi_u u} = \mathbf{R_{u1}}^{-1}\mathbf{Q_{u1}}^{\mathrm{T}}\mathbf{c} - \mathbf{R_{u1}}^{-1}\mathbf{Q_{u1}}^{\mathrm{T}}\mathbf{A_x x}\,. \tag{1.14}$$

This formulation is only valid when each unmeasured variable is observable, that is verified when the columns of matrix $\mathbf{A_u}$ are linearly independent (Narasimhan and Jordache, 2000). In case there are unobservable the factorisation must be properly modified (Loyola-Fuentes and Smith, 2020). This method provides also a classification for the measured variables as redundant or non-redundant, through the inspection of the reduced constraint matrix $\mathbf{Q_{u2}}^{\mathrm{T}}\mathbf{A_x}$. When all the elements in a column of this matrix are zero, then the corresponding measured variable is non-redundant. On the contrary, when a column contains non-zero elements, the corresponding measured variables is classified as redundant, and its value will be adjusted during reconciliation (Loyola-Fuentes and Smith, 2020).

The reduced problem and the consequent QR factorisation can be solved also for a dynamic system (Albuquerque and Biegler, 1986).

1.1.2 Steady-state formulation

In a steady state scenario, the solution of the optimization problem for DR can be calculated using a special case of Equation (1.7) (Ozyurt and Pike, 2004):

$$\hat{\mathbf{x}} = \min_{x_i} \sum_{j=1}^{J} \sum_{i=1}^{I} \left(\frac{y_{ij} - x_i}{\sigma_i}\right)^2 = \min_{x_i} \sum_{j=1}^{J} \sum_{i=1}^{I} {r_{ij}}^2 \tag{1.15}$$

such that

$$c(\hat{\mathbf{x}}, \hat{\mathbf{u}}) = \mathbf{c},$$

$$g(\hat{\mathbf{x}}, \hat{\mathbf{u}}) \geq 0\,,$$

$$Lb_{x_i} \leq \hat{x}_i \leq Ub_{x_i}\,,$$

$$Lb_{u_z} \leq \hat{u}_z \leq Ub_{u_z}\,,$$

where y_{ij} is the measurement j for variable i, x_i is the variable i estimate, σ_i is the standard deviation of the set of measurements $\mathbf{y}_i^{\mathrm{T}}$ and r_{ij} is the respective standardized residual. $\hat{\mathbf{x}}^{\boldsymbol{T}}$ and $\hat{\mathbf{u}}^{\boldsymbol{T}}$ are the $[1 \times I]$ and $[1 \times Z]$ vectors of the final estimates respectively for the measurable variables and unmeasurable variables and c is the process constraint equations. Finally, g represents the inequality constraints while Lb and Ub are the lower and upper bounds for both measured and unmeasured variables.

In a steady-state system, temporal redundancy is exploited because the measurements of the process data are made continuously or discretely in time at a specific sampling rate (Miao et al., 2009). The solution of the optimization problem provides a reconciled value $\hat{x}_i$ for each of the I measured variables will be obtained. As mentioned earlier, in some situations a single snapshot of the process can be considered, leading to the following formulation without temporal redundancy

$$\hat{\mathbf{x}}_j = \min_{x_i} \sum_{i=1}^{I} \left(\frac{y_{ij}-x_i}{\sigma_i}\right)^2 = \min_{x_i} \sum_{i=1}^{I} {r_{ij}}^2, \tag{1.16}$$

subject to the same constraints as in Equation (1.8). At the end vector $\hat{\mathbf{x}}_j$ is obtained, which contains the reconciled values for each of the I variables at specific snapshot j. In this way, if measurements of the considered set of variables are available at several time instants, reconciled values (*i.e.*, steady state values) will be obtained at each sample time. This strategy allows to take into account a possible change in the system steady state due to unwanted effects. On the other hand, it loses accuracy because temporal redundancy which is not exploited to mitigate the effect of bad measurements better (Llanos et al., 2017).

1.1.3 Dynamic formulation

In the case of dynamic systems, the data-time horizon has to consider the system dynamics. The general DR formulation changes to the following

$$\hat{\mathbf{X}} = \min_{\mathbf{x}_t} \sum_{t=0}^{T} \left(\frac{\mathbf{y}_t-\mathbf{x}_t}{\boldsymbol{\sigma}}\right)^2 = \min_{\mathbf{x}_t} \sum_{t=0}^{T} {\mathbf{r}_t}^2 \tag{1.17}$$

such that

$$c\left(\frac{d\hat{\mathbf{x}}(t)}{dt}, \hat{\mathbf{x}}(t), \frac{d\hat{\mathbf{u}}(t)}{dt}, \hat{\mathbf{u}}(t)\right) = \mathbf{c},$$

$$h(\hat{\mathbf{x}}(t), \hat{\mathbf{u}}(t)) = 0,$$

$$g(\hat{\mathbf{x}}(t), \hat{\mathbf{u}}(t)) \geq 0,$$

$$Lb_{x_i} \leq \hat{x}_i(t) \leq Ub_{x_i},$$

$$Lb_{u_z} \leq \hat{u}_z(t) \leq Ub_{u_z} \quad ,$$

where $\mathbf{y}_t^T$, $\mathbf{x}_t{}^T$ and $\mathbf{r}_t{}^T$ are respectively the $[1 \times I]$ measurement, estimate and standardized residual vectors at time instant $t = 0,1, \dots, T$ of the I variables. As before, $\mathbf{u_t}^T$ is the $[1 \times Z]$ vector at time instant t and $\boldsymbol{\sigma}^T$ is the $[1 \times I]$ vector of the variables' standard deviation. $\widehat{\mathbf{X}}$ is the reconciled values matrix $[T \times I]$ which stores the reconciled values for each variable at each time instant. $\mathbf{c}$ is the differential process constraint equations, while h and g are respectively the algebraic and inequality constraints.

1.2 Robust estimators

1.2.1 General characteristics

The previous DR formulations use a Weighted Least Squares (WLS) formulation as the objective function to minimize. However, they are reliable for measurement errors which follow a distribution with zero mean and known variance. For any possible deviation from this assumption, a more appropriate formulation is used, mainly the case when the measurements contain some gross errors (Ozyurt and Pike, 2004). For example, a measurement of a variable affected by gross error can deteriorate the final estimates of the other variables through the DR constraint equations, which relate all system variables. In this way the gross error effect is "smeared" across all of them (Narasimhan and Jordache, 1999). A class of objective functions which are robust to gross errors is that of M-estimators, which are a generalization of the maximum-likelihood estimators (Albuquerque and Biegler, 1996). M-estimators are built in such a way as to concentrate the gross error effect only on the affected variable (Fuente et al., 2015) avoiding the smearing effect.

Their objective function is a loss function ρ which depends on the standardized residual and its shape is specific for the estimator considered. For example, for the WLS, ρ corresponds to the squared standardized residual. Considering this new formulation, Equations (1.15), (1.16) and (1.17) can be rewritten as:

$$\hat{\mathbf{x}} = \min_{x_i} \sum_{j=1}^{J} \sum_{i=1}^{I} \rho\left(\frac{y_{ij} - x_i}{\sigma_i}\right) = \min_{x_i} \sum_{j=1}^{J} \sum_{i=1}^{I} \rho(r_{ij}), \tag{1.18}$$

$$\hat{\mathbf{x}}_j = \min_{x_i} \sum_{i=1}^{I} \rho\left(\frac{y_{ij} - x_i}{\sigma_i}\right) = \min_{x_i} \sum_{i=1}^{I} \rho(r_{ij}), \tag{1.19}$$

$$\widehat{\mathbf{X}} = \min_{\mathbf{x_t}} \sum_{t=0}^{T} \rho\left(\frac{\mathbf{y_t}-\mathbf{x_t}}{\boldsymbol{\sigma}}\right) = \min_{\mathbf{x_t}} \sum_{t=1}^{T} \rho(\mathbf{r}_t), \tag{1.20}$$

maintaining the correspondent set of constraint equations as before

1.2.2 The adopted robust estimators

Simple Method (SiM), Sophisticated Method (SoM), Welsch (W), Quasi-Weighted Least Squares (QWLS), Correntropy (Co) and Fair (F) are the robust DR strategies considered in this Thesis. Their formulation and characteristic are briefly reviewed in the following[1].

1.2.2.1 SiM and SoM

They are both non-adaptive techniques based on the combination of two different estimators (Llanos et al., 2015) in order to reduce the bad effect of outliers on the final estimation. Their formulation consists of two steps for SiM and of an additional third step for SoM:

- Step 1: an estimate of the median of the J measurements of variable i , $\tilde{y}_i$, is calculated using an estimator of the Biweight (BW) family and it corresponds to the solution of the following

$$\tilde{y}_i = \min_{y_i} \sum_{j=1}^{J} \rho_{BW}\left(\frac{y_{ij}-y_i}{\sigma_i}\right), \tag{1.21}$$

  where:

$$\rho_{BW} = \begin{cases} 1-\left[1-\left(r_{ij}/c_{BW}\right)^2\right]^3 \;\; if\;\left|r_{ij}\right| \le c_{BW} \\ 1 \;\; if\left|r_{ij}\right| \ge c_{BW} \end{cases}. \tag{1.22}$$

- Step 2: the solution $\hat{\mathbf{x}}$ of Equation (1.18) is achieved by solving

$$\min_{x_i} \sum_{i=1}^{n} \rho_H\left(\frac{\tilde{y}_i-x_i}{\sigma_i}\right), \tag{1.23}$$

  where:

$$\rho_H = \begin{cases} {r_i}^2 \;\; if\;|r_i| \le c_H \\ 2c_H|r_i| - c_H^2 \;\; if|r_i| \ge c_H \end{cases}, \tag{1.24}$$

  which is a loss function of the Huber (H) family.

1 Note that the following formulation considers the steady-state case exploiting time redundancy. The extension to the non-time redundancy and dynamic case is straightforward and not reported here for the sake of conciseness.

- Step 3: SoM adds an extra third step, which is a refinement of the solution obtained from the second step. The reconciled values obtained from Equation (1.24) are used as initialization points for the following

$$\min_{x_i} \sum_{j=1}^{J} \sum_{i=1}^{n} \rho_{BW} \left(\frac{y_i - x_i}{\sigma_i} \right). \quad (1.25)$$

When temporal redundancy is not used the first step is neglected, losing all its benefits. Therefore, SiM and SoM must be explicitly applied to data within a predefined time window. Furthermore, these two approaches are not designed for dynamic DR, since at the end of Step 1 just one estimate for each variable is produced. This is in conflict with the dynamic behaviour of the system which implies an estimate for each time instant for the same variable.

1.2.2.2 Welsch estimator

The Welsch estimator (Dennis and Welsch, 1976) is usually applied to reconcile measurements of dynamic systems for time varying moving windows (Prata et al., 2008, 2010). Its loss function has the following formulation:

$$\rho_W = c_W^2 \left\{ 1 - exp \left[- \left(\frac{r_{ij}}{c_W} \right)^2 \right] \right\} \quad . \quad (1.26)$$

1.2.2.3 Quasi-Weighted Least Squares estimator

The Quasi-Weighted Least Squares loss function (Zhang et al., 2010) is defined as below

$$\rho_{QWLS} = \frac{{r_{ij}}^2}{2 + c_{QWLS} |r_{ij}|} \quad . \quad (1.27)$$

1.2.2.4 Correntropy estimator

The correntropy estimation loss function is shown in Equation (1.26) and it is a Gaussian kernel function (Chen et al., 2013):

$$\rho_{CO} = -\frac{1}{c_{CO}\sqrt{2\pi}} exp \left[- \left(\frac{{r_{ij}}^2}{2 c_{CO}^2} \right) \right] \quad . \quad (1.28)$$

1.2.2.5 Fair estimator

The Fair estimator (Fair, 1974) behaves like a least square estimator for small residuals, while as an absolute value estimator for large residuals. Its loss function is the following:

$$\rho_F = 2 c_F^2 \left[\frac{|r_{ij}|}{c_F} - ln \left(1 + \frac{|r_{ij}|}{c_F} \right) \right]. \quad (1.29)$$

In particular, for large residuals it behaves as an absolute value estimator. On the other hand, for small residuals it behaves as WLS.

1.2.2.6 Estimators' tuning parameters, efficiency and robustness

Tuning parameters (c_{BW}, c_H, c_W, c_{QWLS}, c_{CO} and c_F) are present in all the loss functions (Ozyurt and Pike, 2004). Depending on their values the loss function changes, influencing the efficiency and robustness of the correspondent estimator.

To understand how well the method performs when the measurement errors follow a known reference distribution, the efficiency is defined according to what reported in Albuquerque and Biegler (1996):

$$E = \frac{V_{opt}}{V_{act}} \tag{1.30}$$

V_{opt} is the optimal error variance (after DR) of an estimator with the reference residual distribution (*i.e.*, standard normal distribution $N(0,1)$), while V_{act} is the residual variance attained by the considered estimator. From the literature (Ozyurt and Pike, 2004; Prata et al., 2010; Llanos et al., 2015), robust M-estimators are usually tuned to reach 95% efficiency relatively to standard normal distribution. For decreasing values of the tuning parameter, the efficiency decreases (Rey, 1983). Furthermore, it is necessary that different estimators have the same efficiency in order to compare their DR performances. In this Thesis estimators were tuned to reach that efficiency and the correspondent tuning parameters values, retrieved from the literature (Rey, 1983; Llanos et al., 2015), are reported in Table 1.1.

Table 1.1. *Tuning parameters for the adopted robust estimators for 95% efficiency relatively to standard normal distribution of the residuals.*

c_{BW}	c_H	c_W	c_{QWLS}	c_{CO}	c_F
4.680	1.345	2.980	0.890	2.050	1.3998

Robustness refers to the characteristic which makes an estimator insensitive to gross deviations from ideality. As for efficiency, tuning parameters influence the robustness of an estimator, where a lower tuning parameter value corresponds to higher robustness (Rey, 1983). In particular, estimator robustness can be assessed by evaluating the influence function. The influence function "*describes the effect of an infinitesimal contamination at the point x on the estimate"* (Hampel et al., 1986) or in other words, the importance of an observation on the estimator. For M-estimators, it is proportional to the first derivative of ρ with respect to the residual (Albuquerque and Biegler, 1986).

Depending on the shape of both loss and influence functions, robust estimators can be classified into two categories: monotone and redescending.

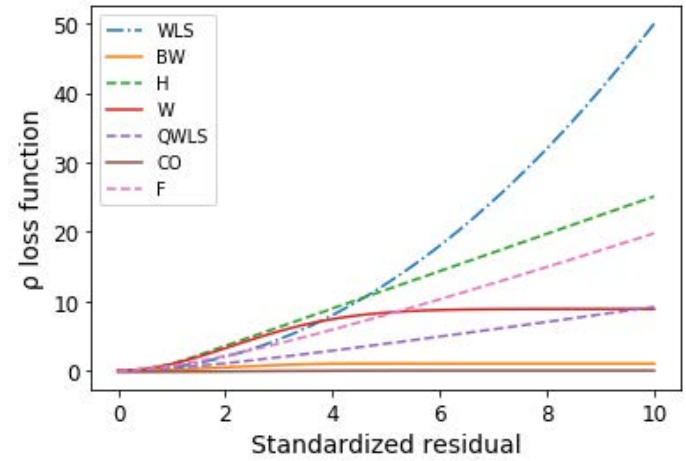
For what concerns monotone estimators (Huber, Fair and QWLS) the influence function reaches a constant value for large residual values, while their loss functions increase in a monotone way and are convex. This determines that when adopting these estimators for DR at the end there will be a unique local minimum. Therefore, the values used to start the iterative solution procedure may influence the number of iterations, but not the final outcome (Llanos et al., 2017).

For what concerns redescending estimators (Welsch, Correntropy and Biweight) the influence function reaches a maximum for low residuals and then their profile "descend" quite fast as residuals' magnitude increases. These estimators are further divided into two sub-classes: *i*) unbounded because at infinite residual magnitude their influence functions tend to zero, or *ii*) bounded when they exactly go to zero. In general, thanks to these characteristics the redescending estimators are best suited to deal with large errors (heavy-tailed error distributions) with respect to the monotones (Llanos et al., 2017). However, they may have several local minima which requires good initialization to ensure attaining an optimal solution.

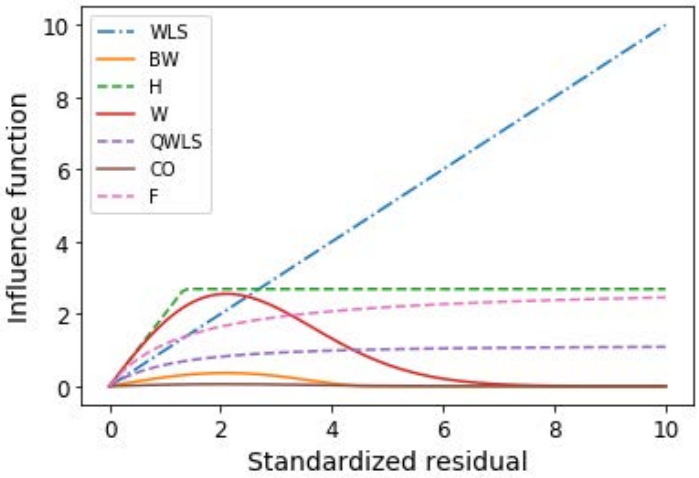
The profiles of the loss and influence function for all the chosen estimators and the Weighted Least Squares are shown in Figures 1.2a and 1.2b. As residuals increase the blue line of WLS influence function (Figure 1.2b) continues to grow and these large errors have a greater effect on the final estimate respect to the smallest ones; this is why WLS is not robust. On the contrary the influence function of the other estimators remains bounded, allowing them to be robust to increasing errors. Furthermore, two different behaviours can be noticed:

- the dashed lines of Huber, Fair and QWLS influence functions reach a constant value for large residual values, while their loss functions increase. They are monotone and convex estimators[2];
- the solid lines for Welsch, Correntropy and Biweight highlight their redescending behaviour. In particular, the first ones have unbounded ρ because at infinite residual magnitude their influence functions tend to zero, while Biweight has bounded ρ because its influence function goes exactly to zero after a certain threshold.

2 Their convexity cannot be appreciated from the figure because only positive residuals are considered in Figure 1.2.

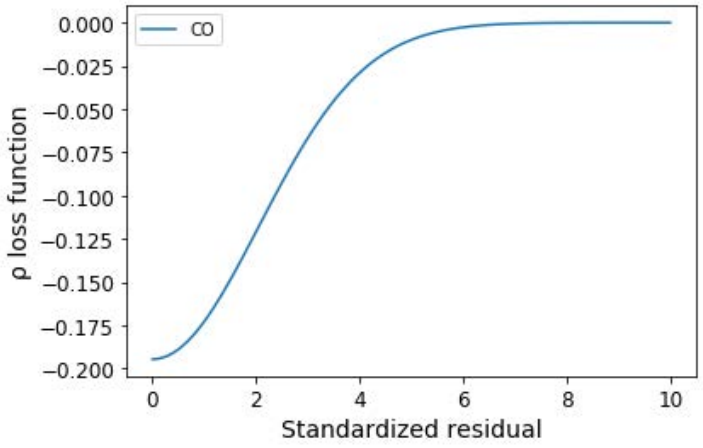


(a)

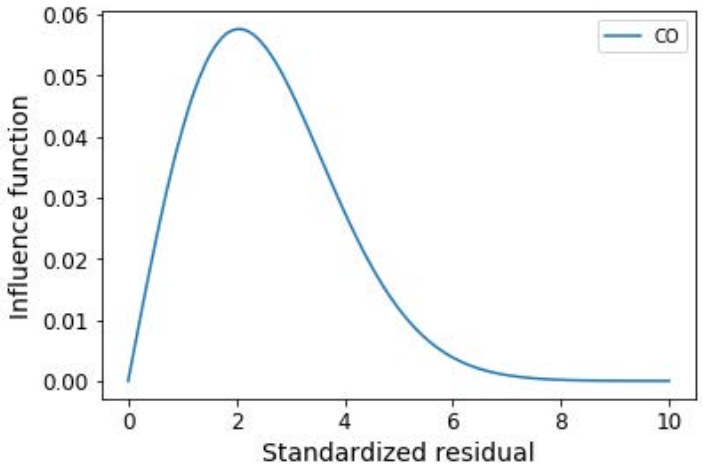


(b)

Figure 1.2. *Loss function (a) and influence function (b) for the adopted estimators tuned at 95% efficiency and for the Weighted Least Squares (WLS).*



(a)



(b)

Figure 1.3. *Detailed view of loss function (a) and influence function (b) for Correntropy redescending estimator tuned at 95% efficiency*

1.3 Gross error detection techniques

Gross error detection (GED) is a technique that has been developed to detect the presence of gross errors in the measurements dataset and identify the affected variables. In general, two different situations may rise depending on the quantity of variables affected by gross errors. When one variable contains gross errors, a single GED must be performed; when more than one variable contains gross errors a multiple GED (MGED) is necessary and the affected variables can be detected one by one or simultaneously (Romagnoli and Sanchez, 2000). In

any case, DR and GED are applied together to improve the accuracy of the final estimates. In the past decades (Romagnoli and Sanchez, 2000; Miao et al., 2009) many GED approaches have been developed and analysed. In this Thesis two approaches are used, which are suitable to be implemented with robust estimators (Ozyurt and Pike, 2004): "cut-off points" and X84.

1.3.1 Cut-off points

The influence function of a robust estimator can be exploited in order to simultaneously perform DR and GED, by using a series of cut-off points $cp_{estimator}$ as a discriminant for the residuals' magnitude after DR. This method does not imply the knowledge of the prior measurements' distribution (Ozyurt and Pike, 2004). According to this technique a gross error is therefore identified if

$$\left|r_{ij,R}\right| \geq cp_{estimator} \tag{1.31}$$

where

$$r_{ij,R} = \frac{y_{ij} - \hat{x}_i}{\sigma_i} \tag{1.32}$$

is the standardized residual after DR for measurement j of variable i[3]. Therefore, if the condition (1.29) is verified an outlier is detected; in case of four outliers are found out of five consecutive observations a bias is identified (Prata et al., 2010).

At least, four different $cp_{estimator}$ values can be assigned as criterion: the maximum of the influence function, its inflection point and two inflection points of the first derivative of the influence function. In this Thesis just the maxima and the second inflection point of the first derivative of the influence function were considered because they were proven to grant the best GED results (Ozyurt and Pike, 2004). From Figure 1 it is clear that only redescending estimators have cut-points contrarily to the monotone ones. However, provided that the functions have the same efficiencies, the cut points for redescending influence functions can be considered a cut point candidate for the non-redescending estimators (Ozyurt and Pike, 2004).

Cut-off points GED can be applied to both steady-state and dynamic systems (Rey, 1983) and to residuals of a fixed horizon time window or of a moving time window. It is suitable for

3 Note that in this case $r_{ij,R}$ is referred to steady-state DR exploiting time redundancy. For steady non-time redundancy DR and dynamic DR, a different reconciled value $\hat{x}_i$ must be considered in accordance.

simultaneously performing DR and simultaneous detection/identification of multiple gross errors by exploiting the same influence function for two different tasks.

1.3.2 X84

X84 (Hampel et al., 1986) rejection criterion is considered in this work as an alternative to the previous GED algorithms. This technique rejects the measurements whose residuals after DR are more than 5.2 (commonly assumed value) times the median deviation away from the median of the residuals. In this case, the median deviation (median absolute deviation) corresponds to the median of the absolute value of the difference between the residuals and their median. This rule can be applied to any ρ-function, not relying on a predetermined and ρ-function-dependent cut-off point.

CHAPTER 2

Case studies

In this Chapter three case studies of heat exchange in the refinery sector are illustrated: two simulated cases (Singhmaneeskulchai, Siemanond et al., 2013; Kongchuay and Siemanond, 2014) (namely, one steady-state case and one dynamic), and a steady-state industrial example (Coletti and Macchietto, 2011).

2.1 Simulated steady-state case study

2.1.1 Simulated system

The simulated steady-state case study considered in this Thesis is a shell-and-tube heat exchanger presented in the literature by Kongchuay and Siemanond (2014). The involved fluids are a dowthermal oil, as a hot fluid, and ethane, as cold fluid to heat-up. The following six process variables were measured: the inlet temperature of the oil $T_{o,in}$, the oulet temperature of the oil $T_{o,out}$,the inlet temperature of the ethane $T_{e,in}$, ,the outlet temperature of the ethane $T_{e,out}$, the oil volumetric flowrate $\dot{V}_o$ and the ethane volumetric flowrate $\dot{V}_e$.

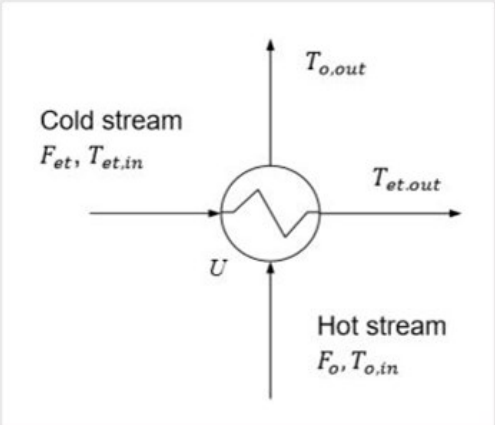


Figure 2.1. *Simulated single shell-and-tube heat exchanger with the six measured process variables. The hot stream is a diathermic oil, while the cold one is ethane (Kongchuay and Siemanond, 2014).*

The system is assumed at its steady state. The material balance is assumed to be satisfied for both fluids and it is explicitly embedded in the two energy balances at the hot and cold side, respectively:

$$-Q = \dot{m}_o C_{p,o}\left(T_{o,out} - T_{o,in}\right) \tag{2.1}$$

$$Q = \dot{m}_e C_{p,e}\left(T_{e,out} - T_{e,in}\right) \tag{2.2}$$

with

$$Q = UA\Delta T_{ml}\,, \tag{2.3}$$

where Q is heat-duty of the heat exchanger in [W] and UA is the "thermal size" in [W/°C] composed respectively by the overall heat transfer coefficient *U* and the exchange area *A*. $\dot{m}_o$ and $\dot{m}_e$ are the fluids mass flowrates in [kg/s], while $C_{p,o}$ and $C_{p,e}$ are the respective specific heats in [J/kg °C]. Finally, ΔT_{ml} is the mean logarithmic temperature calculated according to the approximation proposed by Chen (2019) considering the two fluids counter-current:

$$\Delta T_{\mathrm{ml}} = \left[\left(T_{o,in} - T_{e,out}\right)\left(T_{o,out} - T_{e,in}\right)0.5\left(T_{o,in} - T_{e,out} + T_{o,out} - T_{e,in}\right)\right]^{1/3} \tag{2.4}$$

The 4 degrees of freedom of the considered system were saturated by fixing the four temperatures and solving the balance equations in order to determine the two volumetric flowrates. The system of equations was implemented in Python and solved through the `SciPy` library.

These six values correspond to the true values, necessary for measurements generation and DR/GED performance evaluation. In the Tables 2.1 and 2.2 the true values and physical parameters are reported. In particular, the temperatures and parameters (assumed independent from temperature variations) were retrieved from the literature (Singhmaneeskulchai, Siemanond et al., 2013; Kongchuay and Siemanond, 2014).

Table 2.1. *True values of the six measured process variables.*

$T_{h,in}$ [°C]	$T_{h,out}$ [°C]	$T_{e,in}$ [°C]	$T_{e,out}$[°C]	$\dot{V}_o$ [m³/h]	$\dot{V}_e$ [m³/h]
170	103	16	65	39.4	30614.44

Table 2. 2. *Physical parameters for the two fluids and heat-exchanger thermal size.*

$C_{p,e}$[J/kg°C]	ρ_e [kg/m³]	$C_{p,o}$ [J/kg°C]	ρ_o [kg/m³]	UA [W/°C]	$\dot{V}_e$ [m³/h]
2473.00	1.33	2419.00	772.65	14318.66	2473.00

2.1.2 Measurements generation

Six simulated datasets of 365 measurements for each variable were generated according to the scenarios reported below.

1) **Scenario 1**: measurements affected by random noise. The 365 data were randomly generated from a normal distribution with a mean which is equal to the true values in Table 2.1 and the lowest level of standard deviation reported in Table 2.3;
2) **Scenario 2**: measurements affected by random noise with the highest level of standard deviation reported in Table 2.3.
3) **Scenario 3**: measurements affected by low noise and 40 outliers on three variables ($T_{o,in}$, $T_{e,in}$, $\dot{V}_o$): 20 were generated by subtracting a value between 15% and 150% of the true value from the correspondent measurement; the other 20 were obtained by adding a value between 15% and 150% of the true value instead.
4) **Scenario 4**: measurements affected by low noise and 120 outliers on variable $T_{e,in}$: 60 were generated by subtracting a value between 50% and 200% of the true value from the correspondent measurement; the other 60 were obtained by adding a value between 50% and 200% of the true value instead.
5) **Scenario 5:** measurements affected by low noise and biases on variables $T_{e,in}$, $\dot{V}_o$, adding 5 [°C] and 10 [m^3/h] to the measurements, respectively.
6) **Scenario 6**: measurements affected by low noise and biases on variables $T_{e,in}$, $\dot{V}_o$, adding 10 [°C] and 20 [m^3/h] to the measurements, respectively.

Table 2.3. *Process variables standard deviations used for noise generation in Scenario 1 and Scenario 2.*

noise	$T_{h,in}$ *[°C]*	$T_{h,out}$ *[°C]*	$T_{e,in}$ *[°C]*	$T_{e,out}$ *[°C]*	$\dot{V}_o$ *[m^3/h]*	$\dot{V}_e$ *[m^3/h]*
low	5	5	2	2	2	5
high	25	25	10	10	10	25

In this way, a variety of scenarios were tested, allowing to evaluate the adopted estimators in a realistic way. All the scenarios were generated through the `numpy.random` command in Python. The values for the standard deviation of the normal distributions for Scenario 2 were retrieved from the literature (Kongchuay and Siemanond, 2014), and they were lowered for

Scenario 1. The magnitude of the outliers and biases were also chosen considering other simulated cases performed by other authors (Ozyurt and Pike, 2004; Llanos et al., 2017).

2.2 Simulated dynamic case study

2.2.1 Simulated system

The same single shell-and-tube heat exchanger of the previous case was simulated as a dynamic system (Singhmaneeskulchai, Siemanond et al., 2013). The two fluids remained unchanged, as well as their physical properties. The outlet temperatures ($T_{o,out}$, $T_{e,out}$) are assumed at steady-state, while the remaining ones ($T_{o,in}$, $T_{e,in}$, $\dot{V}_o$, $\dot{V}_e$) have a dynamic behaviour. The degrees of freedom are saturated by fixing the two outlet temperatures and the parameters needed to determine the dynamics of the inlet temperatures, which are perturbed through a sinusoidal change:

$$T_{o,in}(t) = T_{o,in}(0) + A_o \sin(\omega_o t) \tag{2.5}$$

$$T_{e,in}(t) = T_{e,in}(0) + A_e \sin(\omega_e t)\,. \tag{2.6}$$

A_o and A_e are the two oscillation-amplitudes respectively equal to 10 [°C] and 5 [°C], while ω_o and ω_e are the angular frequencies both set to 1 [rad/s]. These values are chosen because they yield the most satisfactory time-profile. The system differential energy equations are presented by:

$$\frac{dq}{dt} = \left(\frac{rho_e\, cp_e}{3600}\right)\left[T_{e,out}\frac{dV_e}{dt} - T_{e,in}\frac{dV_e}{dt} - V_e\frac{\mathrm{dT_{e,in}}}{dt}\right] \tag{2.7}$$

$$\frac{dq}{dt} = -\left(\frac{rho_o\, cp_o}{3600}\right)\left[T_{o,out}\frac{dV_o}{dt} - T_{o,in}\frac{dV_o}{dt} - V_o\frac{\mathrm{dT_{o,in}}}{dt}\right] \tag{2.8}$$

where,

$$\frac{dq}{dt} = UA\frac{d\Delta T_{ml}}{dt}\,. \tag{2.9}$$

As in the steady-state case, the material balance is assumed to be satisfied for both fluids and it is explicitly embedded in the energy balances above.

The solution of these equations yielded the time-profiles of the two volumetric flowrates which, together with the profiles of the temperatures, correspond to the true values time-profiles. As in the steady-state case, these values were necessary for the measurements'

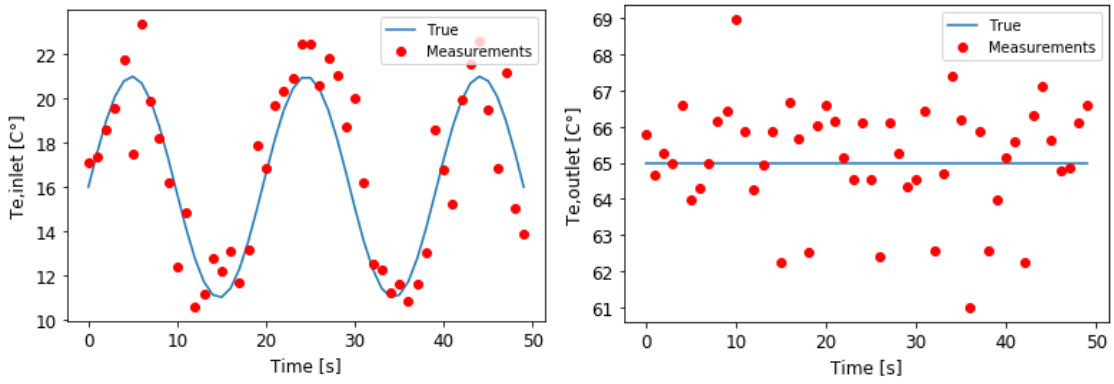
generation and DR/GED performance evaluation. As initial guesses for the flowrates the correspondent steady-state values, reported in Table 2.1, were used; in the same way $T_{o,in}(0)$, $T_{e,in}(0)$, $T_{o,out}$ and $T_{e,out}$ were all set equal to the previous steady values.

The system was implemented in Pyhton and the system of differential equations was solved by applying the Forward Euler method. This choice is driven by the explicit behaviour of this method and its straightforward implementation.

2.2.2 Measurements generation

Six different data scenarios were generated once again by using `numpy.random`, and measurements were generated at time steps of 1 s for 50 s for all the variable.

1) **Scenario 1**: measurements affected by low random noise. A single normal distribution was generated for each time instant (for each dynamic variable) with mean equal to the respective true value (at that time instant) and low standard deviation, as reported in Table 2.3. Two examples of measurements and true values for a dynamic and a steady variable are shown in Figure 2.2;



(a) (b)

Figure 2.2. *Measurements of Scenario 1 and true values for ethane inlet temperature (a) and ethane outlet temperature (b).*

2) **Scenario 2**: measurements affected by high random noise;
3) **Scenario 3**: measurements affected by low noise and 5 outliers on three variables ($T_{o,in}$, $T_{e,in}$, $\dot{V}_o$), generated by adding a value between 100% and 150% of the true value from the correspondent measurement.

4) **Scenario 4**: measurements affected by low noise and 5 outliers on variable $T_{e,in}$, generated by adding a value between 200% and 250%.
5) **Scenario 5:** measurements affected by low noise and biases on variables $T_{e,in}$, $\dot{V}_o$: 20 consecutive data affected by a bias of 5 [°C] and 10 [m^3/h] respectively.
6) **Scenario 6**: measurements affected by low noise and biases on variables $T_{e,in}$, $\dot{V}_o$: 20 consecutive data affected by bias of 10 [°C] and 20 [m^3/h], respectively.

2.3 Industrial case study

2.3.1 Industrial system

The industrial case study considered in this Thesis is based on a network of heat exchangers originally presented by Coletti and Macchietto (2011) whose simplified flowsheet is shown in Figure 2.3. It represents the hot end of a pre-heat train for crude oil before the crude distillation column.

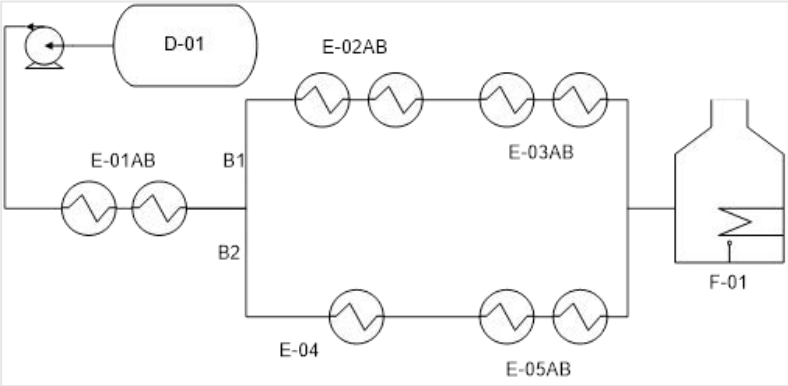


Figure 2.3. *Simplified flowsheet of the hot end of a pre-heat train for crude oil (Coletti and Macchietto, 2011).*

The considered heat exchanger network (HEN) begins downstream of the pre-flash drum (D-01) where a crude oil passes through 5 shell-and-tube heat exchanger units before the furnace (F-01), where it is heated-up to 340 [°C]. The crude flows on the tube-side in all units, while on the shell-side different hot crude fractions from the distillation column flow. The API degree and the coded name of the involved fluids are reported in Table 2.4. Downstream the first unit (heat exchanger E-01AB), the crude stream splits in two branches B1 and B2, and re-connects before the furnace through a mixer. The branch, B1, comprises heat exchanging units E-02 and E-03 while the branch B2, comprises units E-04 and E-05.

Table 2.4. *Coded HEN fluids and their relative API degree.*

Fluid	*Crude*	*SS1*	*SS2*	*SS3*	*SS4*
API	*35.00*	*26.37*	*17.09*	*18.44*	*29.84*

In Table 2.5 the main geometrical parameters of the heat exchangers are reported. It can be seen that all five heat exchangers have counter-current inlets, E-04 has a single shell and the others are double shell.

Table 2.5. *Geometrical parameters for the five heat exchangers of the considered network. The term "c.c" is for counter-current arrangement.*

Parameters	*E-01AB*	*E-02AB*	*E-03AB*	*E-04*	*E-05AB*
Arrang.	c.c.	c.c.	c.c.	c.c.	c.c.
No. shells	2	2	2	1	2
No. tube pass	4	8	4	4	8

2.3.2 Industrial measurements

This network comprises 5 streams for which temperature and flowrate measurements are available for five years of operations. The correspondent nominal values are not available. These variables will be denoted through the names of the correspondent sensors. For flowrates it will be FC (flow controllers) followed by a progressive number (*i.e.*, FC001…, FC007); for temperatures it will be TI (temperature indicator) followed by a progressive number (*i.e.*, TI001…, TI017) instead. In particular, the temperature measurements are related to dirt heat exchangers.

6.3% of the dataset elements are found to be Not A Number (NaN), due to recording errors, and the correspondent set of measurements are deleted. The behaviour of these datapoints can be sum up by their standard deviations, as reported in Table 2.6.

The positions of the sensors led to the presence of unmeasured variables in some streams and therefore along with the measured variables there were also unmeasured flowrates and temperatures.

Table 2.6. *Standard deviations of the measured flowrates (a) and temperatures (b).*

(a)

Variable	*σ [kg/s]*
FC001	14.46
FC002	12.04
FC003	13.50
FC004	5.19
FC005	10.28
FC006	4.33
FC007	11.58

(b)

Variable	*σ [°C]*
TI001	7.10
TI002	9.64
TI003	9.61
TI004	8.30
TI005	10.17
TI006	10.58
TI007	10.43
TI008	11.41
TI009	4.84
TI010	30.01
TI011	9.15
TI012	26.20
TI013	21.06
TI014	9.62
TI015	24.59
TI016	14.06
TI017	6.66

2.3.3 Unit models

The HEN was modelled according to the following assumptions:

- the system is at steady-state within the considered time-frame;
- The effectiveness of the heat exchangers was not considered;
- No fouling effect or model was accounted.

The system is described according to the mass and energy balances for each unit:

- heat exchangers: the mass balances on a single heat-exchanger for hot and cold side, respectively, are:

$$m_{h,i} - m_{h,o} = 0 \tag{2.10}$$

$$m_{c,i} - m_{c,o} = 0\,, \tag{2.11}$$

where the subscripts h and c represent the hot and cold streams, and the subscripts i and o represent the heat exchanger inlet and outlet.

The energy balance for the cold and hot side were evaluated by:

$$C_h = m_{h,i}\frac{(Cp_{h,i}+Cp_{h,o})}{2} \tag{2.12}$$

$$C_c = m_{c,i}\frac{(Cp_{c,i}+Cp_{c,o})}{2}, \tag{2.13}$$

where the stream Cps are estimated from the retrieved measurements using engineering judgment. C_h and C_c are the heat capacities of the hot and cold sides, respectively.

The energy balance around the heat-exchanger yields

$$C_c(T_{c,o} - T_{c,i}) = C_h(T_{h,i} - T_{h,o}) \tag{2.14}$$

which corresponds to

$$T_{c,i} - T_{c,o} + C_R T_{h,i} - C_R T_{h,o} = 0\,, \tag{2.15}$$

where C_R is the ratio between the lower and the larger heat capacity flow rates.

If the heat capacity flow rate of the hot stream is larger than heat capacity flow rate of the cold stream, a correspondent equation to Equation (2.15) can be developed:

$$-C_R T_{c,i} + C_R T_{c,o} - T_{h,i} + T_{h,i} = 0\,. \tag{2.16}$$

The set of Equations (2.15) and (2.16) can be organized in a single structure, through the definition of the parameter y, where $y = 1$, if the hot stream has the lower heat capacity flowrate, and $y = 0$, if the cold stream has the lower heat capacity flowrate. Thus,

$$[C_R(y-1)+y \quad -C_R(y-1)-y \quad C_R y+(y-1) \quad -C_R y-(y-1)]\begin{bmatrix} T_{c,i} \\ T_{c,o} \\ T_{h,i} \\ T_{h,o} \end{bmatrix} = \mathbf{0} \tag{2.17}$$

- mixer: several inlet streams generate a unique outlet stream. Mixer model equations are valid for hot and cold streams. Therefore, the mass balance results

$$m_{i,1} + m_{i,2} + m_{i,3} + \cdots - m_o = 0\,, \tag{2.18}$$

and the correspondent energy balance is

$$C_{i,1}T_{i,1} + C_{i,2}T_{i,2} + C_{i,3}T_{i,3} + \cdots - C_o T_o = 0\,. \tag{2.19}$$

- splitter: the mass balance equations have to consider the split fraction parameter α as follows and are valid for hot and cold streams:

$$m_i - m_{o,1} - m_{o,2} = 0 \tag{2.20}$$

$$\alpha m_i - m_{o,1} = 0\,. \tag{2.21}$$

The energy description involves an additional equation which establishes that the temperature between inlet and outlet streams cannot change:

$$C_i T_i - C_{o,1} T_{o,1} - C_{o,2} T_{o,2} = 0\,, \tag{2.22}$$

$$T_i - T_{o,1} = 0\,. \tag{2.23}$$

2.3.4 Heat exchanger network model

According to the network connectivity, the previous individual unit models must be interrelated, building the HEN model. The HEN model is composed by two systems of equations: a network flow model (*i.e.*, mass balance equations) and a network energy model (*i.e.*, energy balance equations) (de Oliveira Filho et al., 2007).

According to this method, the topology of the heat exchanger network is represented by a digraph consisting in edges and vertex (Mah, 1990), representing internal streams and HEN units, respectively. HEN units are denoted by heat exchangers (HE), splitter (SP), mixer (MX) and furnace (FUR). Therefore, the network connectivity can be represented using a $[N \times S]$ incidence matrix **M**, Equation (2.24), which contains the interactions among S edges and N vertices. S can be further divided into S_c cold streams and S_h hot streams. In the same way N is the sum of the total number N^{HE} of heat exchangers, the number N^{MX} of mixers, the number N^{SP} of splitters and the number N^{FUR} of furnaces. The elements of **M** assume values of -1, 0 and 1, depending if a stream is directed from, is not directed or is directed to a unit, respectively, meaning that **M** is made of a series of submatrices:

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}^{HE} \\ \mathbf{M}^{MX} \\ \mathbf{M}^{SP} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_c^{HE} & \mathbf{M}_h^{HE} \\ \mathbf{M}_c^{MX} & \mathbf{M}_h^{MX} \\ \mathbf{M}_c^{SP} & \mathbf{M}_h^{SP} \end{bmatrix}. \tag{2.24}$$

The incidence matrix can be exploited to formulate the two following network models for mass and energy:

$$\mathbf{Ax} = \mathbf{b}\,, \tag{2.25}$$

$$\mathbf{Cz} = \mathbf{d} \, . \tag{2.26}$$

The matrices **A** and **C**, and vectors **b** and **d** comprise of the linear mass and energy balance equations, respectively, which are formulated in terms of incidence matrices **M**. These equations are solved for the vector of network flow rates $\mathbf{x}\ [7 \times 1]$ and the vector of network temperatures $\mathbf{z}\ [17 \times 1]$, respectively. Details regarding the general formalization of the full model in terms of incidence matrices can be found in the work by Loyola-Fuentes (2019) and in de Oliveira Filho et al. (2007).

CHAPTER 3

Results of the data reconciliation and gross error detection for the simulated heat exchanger case study

In this Chapter the results of the two simulated case studies are presented. The methodologies and performance indexes applied for DR and GED are firstly discussed, before illustrating the results.

3.1 Results for the simulated steady-state case study

3.1.1 Evaluation of the data reconciliation and gross error detection performance

DR was performed by adopting the formulation which exploits and does not exploits temporal redundancy, Equations (1.18) and (1.19). The heat-exchanger energy balance equations, Equations (2.1), (2.2) and (2.3) were used as the DR system constraints.

In order to evaluate and compare the performances of DR with different estimators in all the simulated scenarios, two performance indexes were calculated.

The first one is the sum of squared error (SSE),

$$\mathrm{SSE} = \sum_{i=1}^{n} \left(\frac{\widehat{x_i} - x_i}{\sigma_i} \right)^2, \tag{3.1}$$

$$\mathrm{SSE_j} = \sum_{i=1}^{n} \left(\frac{\widehat{x_{ij}} - x_i}{\sigma_i} \right)^2, \tag{3.2}$$

Where $\hat{x}_i$, x_i and σ_i are respectively the reconciled value, the true value and the measurements standard deviation for variable i. Equation (3.1) is applied in case of reconciliation exploiting time-redundancy, while Equation (3.2) is for no-time redundancy, referring to a certain snapshot j .

The other relevant index is total error reduction (TER) (Ozyurt and Pike, 2004) whose Equation is defined as follow:

$$TER = 100\frac{\sqrt{\sum_{i=1}^{n}((y_{ij}-x_i)/\sigma_i)^2}-\sqrt{\sum_{i=1}^{n}((\hat{x}_i-x_i)/\sigma_i)^2}}{\sqrt{\sum_{i=1}^{n}((y_{ij}-x_{ij})/\sigma_i)^2}}\ , \tag{3.3}$$

$$TER_j = 100\frac{\sqrt{\sum_{i=1}^{n}((y_{ij}-x_i)/\sigma_i)^2}-\sqrt{\sum_{i=1}^{n}((\hat{x_{ij}}-x_i)/\sigma_i)^2}}{\sqrt{\sum_{i=1}^{n}((y_{ij}-x_i)/\sigma_i)^2}}\ , \tag{3.4}$$

where y_{ij} is the measured value of variable i at measurement/snapshot j. Equation (3.3) is applied in case of reconciliation exploiting time-redundancy, while Equation (3.4) for non-time redundancy.

It is desirable to have low SSE values and high TER values. This means that the reconciled values are close to the true ones and the measurements errors have been successfully reduced. A negative TER means that the error is increased after DR.

The DR algorithm was implemented in Python and the constrained optimization problem was solved by adopting the Sequential Least Squares Quadratic Programming (SLSQP; Nocedal and Wright, 2006). The calculations are executed using a Processor Intel Core (TM) i7-7500U CPU @ 2.90 GHz, 16GB RAM which is adopted also for the next case studies. The measurements median was used to initialize the problem when adopting the time redundancy formulation, being a robust trend indicator. In the case without redundancy the measurements of the correspondent snapshot were used instead. In fact, the median embeds all the snapshots and it would be inappropriate as initial guess of a problem concerning just a single snapshot.

In order to evaluate the GED performances of each estimator in all the scenarios, the following parameters are calculated: average number of type I error (AVTI) (Ozyurt and Pike, 2004), Equation (3.5) , and overall power (OP) (Ozyurt and Pike, 2004), Equation (3.6), respectively:

$$AVTI = \frac{Number\ of\ \ wrongly\ identified\ gross\ errors}{Number\ of\ simulations}\ , \tag{3.5}$$

$$OP = \frac{Number\ of\ correctly\ identified\ gross\ errros}{Total\ number\ of\ gross\ errors}. \tag{3.6}$$

In AVTI the numerator indicates the number of measurements which do not contain gross errors but they are detected as containing instead; its denominator refers to the number of simulations for the same scenario, in this Thesis just one simulation for each scenario was performed. Regarding OP, the denominator considers the measurements which are really affected by the simulated gross errors, while the denominator refers to the number of

measurements where a gross error is present (*i.e.*, number of simulated outliers or number of measurements affected by bias). The optimal GED technique should lead to low AVTI value and high OP value, detecting all and only the true gross errors.

Table 3.1. *Cut-off points adopted for GED for the seven estimators adopted in this Thesis.*

Estimator	***Cut-points***
Simple Method	Biweight influence function: 2.093 (max.), 2.093 (i.p.)
Sophisticated Method	Biweight influence function: 2.093 (max.), 2.093 (i.p.)
Welsch	Welsch influence function: 2.11 (max.), 4.92 (i.p.)
QWLS	4.51 (α=0.05) and 4.78 (α=0.025) from $N(0,1)$
Correntropy	Correntropy influence function: 2.05 (max.), 4.78 (i.p.)
Fair	Contaminated Normal influence function: 2.13 (max.), 3.34 (i.p.)
WLS	4.51 (α=0.05) and 4.78 (α=0.025) from $N(0,1)$

3.1.2 Results

3.1.2.1 Data reconciliation

In Tables 3.2a and 3.2b the SSE and the median SSE are reported for all the tested estimators and all the scenarios for both DR formulations. The reconciled values are much closer (lower SSE) to the true ones when temporal redundancy is available. This is due to the beneficial effect of exploiting all available measurements which mitigates the effect of outliers, improving the quality of the final estimates.

Comparing in the specific the estimators, the following observations can be done:

- SiM and SoM yield the best reconciled values in the time redundancy case, taking advantage from the reconciliation provided by the first Step. However, they lose their superiority in the last two Scenarios;
- in the case without time redundancy, the median SSE is substantially the same for each estimator in the same Scenario. However, as bias increases Welsch and Correntropy have a clear superiority;

From these results it is highlighted that the presence of bias is an issue for robust estimators as already highlighted by Prata et al. (2008). The use of temporal redundancy can mitigate this negative effect but the presence of more reconciliation steps can emphasize it, as happened for SiM and SoM. In absence of temporal redundancy, the greater robustness of redescending estimators, Welsch and Correntropy, proves to be a good protection.

Table 3.2. *Simulated steady-state case study: SSE for the time-redundancy DR (a) and median SSE for the non-time redundancy DR (b) of each estimator and scenario.*

(a)

SSE with time redundancy							
Scenario	**SiM**	**SoM**	**Welsch**	**QWLS**	**Correntropy**	**Fair**	**WLS**
1	$2{\cdot}10^{-3}$	$2{\cdot}10^{-3}$	$1.4{\cdot}10^{-2}$	$1.5{\cdot}10^{-2}$	$1.4{\cdot}10^{-2}$	$1.5{\cdot}10^{-2}$	$1.4{\cdot}10^{-2}$
2	$2{\cdot}10^{-3}$	$2{\cdot}10^{-3}$	$1.4{\cdot}10^{-2}$	$1.5{\cdot}10^{-2}$	$1.5{\cdot}10^{-2}$	$1.5{\cdot}10^{-2}$	$1.4{\cdot}10^{-2}$
3	$2{\cdot}10^{-3}$	$1\cdot10^{-3}$	$2{\cdot}10^{-3}$	$2{\cdot}10^{-3}$	$2{\cdot}10^{-3}$	$2{\cdot}10^{-3}$	$2{\cdot}10^{-3}$
4	$6\cdot10^{-3}$	$6\cdot10^{-3}$	$1.6{\cdot}10^{-2}$	$1.6{\cdot}10^{-2}$	$1.6{\cdot}10^{-2}$	$1.6{\cdot}10^{-2}$	$1.7{\cdot}10^{-2}$
5	29.789	29.789	8.934	8.942	8.933	8.934	8.935
6	143.345	143.345	45.734	45.313	45.851	45.677	45.565

(b)

SSE without time redundancy							
Scenario	**SiM**	**SoM**	**Welsch**	**QWLS**	**Correntropy**	**Fair**	**WLS**
1	3.31	3.31	3.32	3.29	3.32	3.29	3.27
2	3.38	3.41	3.42	3.41	3.43	3.41	3.31
3	2.76	2.76	2.89	2.78	2.90	2.79	2.84
4	3.58	3.57	3.65	3.61	3.67	3.60	3.46
5	34.43	38.15	34.16	35.76	33.78	35.61	31.62
6	145.13	173.05	74.94	143.54	65.88	141.77	123.76

An unexpected result is that the DR performances of WLS result comparable to those of robust estimators. According to the theory, it is expected to behave quite well when only noise is present (Scenarios 1 and 2) and in a poor way in presence of outliers/bias (Scenarios 3-6).
This is probably due because the measurements' residuals are not great enough. They do not fall in the region where the WLS influence function begins to diverge from those of robust estimators. In this way, the benefits of having a bounded influence function cannot be exploited. However, this seems to not happen in Scenario 6, where the SSE of WLS is higher than those of Welsch and Correntropy, the two most robust estimators. This is likely due to the smearing effect which affects the WLS final estimates. Under this perspective, a close look to the median squared errors (SE) in Table 3.3 is helpful:

- as highlighted in the first row of Tables 3.3a and 3.3b, in Scenario 5 both estimators cannot lead to good reconciliation for the two variables affected by bias ($T_{e,in}$, $\dot{V}_o$). The other variables have quite low SE values: they are not influenced by the faulty variables;
- for what concerns the second row of Tables 3.3a and 3.3b, in Scenario 6 there is an increase of the SE of $T_{e,in}$ and $\dot{V}_o$. However, for Correntropy the SE of the other variables remains low, contrarily to WLS where there is a strong increase; using this estimator the error due to bias is propagated.

Table 3.3. *Simulated steady-state case study: median squared error (SE) of the six variables in the non-time redundancy for Correntropy (a) and WLS (b) estimators in Scenario 5 and 6.*

(a)

SE Correntropy						
Scenario	$\mathbf{T_{h,in}}$	$\mathbf{T_{h,out}}$	$\mathbf{T_{e,in}}$	$\mathbf{T_{e,out}}$	$\mathbf{\dot{V}_o}$	$\mathbf{\dot{V}_e}$
5	2.31	1.51	4.65	0.85	19.56	0.48
6	1.77	1.07	22.86	1.95	25.53	0.478

(b)

SE WLS						
Scenario	$\mathbf{T_{h,in}}$	$\mathbf{T_{h,out}}$	$\mathbf{T_{e,in}}$	$\mathbf{T_{e,out}}$	$\mathbf{\dot{V}_o}$	$\mathbf{\dot{V}_e}$
5	2.59	1.44	3.36	0.91	19.78	0.48
6	8.88	3.97	14.95	3.30	88.41	0.48

Table 3.4. *Simulated steady-state case study: median TER for the time-redundancy DR (a) and for the non-time redundancy DR (b) of each estimator and scenario.*

(a)

Median TER with time redundancy							
Scenario	**SiM**	**SoM**	**Welsch**	**QWLS**	**Correntropy**	**Fair**	**WLS**
1	97.83	97.84	94.58	94.55	94.57	94.55	94.6
2	97.83	97.84	94.57	94.53	94.57	94.54	94.58
3	98.05	98.05	97.89	97.84	97.89	97.84	97.93
4	96.52	96.52	94.48	94.44	94.48	94.44	94.42
5	10.83	10.83	51.17	51.14	51.17	51.16	51.16
6	-2.67	-2.67	42.01	42.27	41.93	42.04	42.11

(b)

Median TER without time redundancy							
Scenario	**SiM**	**SoM**	**Welsch**	**QWLS**	**Correntropy**	**Fair**	**WLS**
1	15.88	15.26	15.01	15.47	14.95	15.45	15.91
2	15.71	15.44	15.22	14.99	15.16	14.99	16.19
3	4.36	4.2	3.26	3.39	3.26	3.36	4.41
4	12.86	13.01	11.87	12.23	11.83	12.16	12.86
5	3.7	2.17	3.08	1.66	2.93	2.03	7.84
6	-3.29	-13.54	28.64	-2.79	30.73	-2.17	4.35

Similar results can be found studying the TER index (Table 3.4) but with new evidence:

- exploiting temporal redundancy almost completely reduces the error, except when bias is present;
- the presence of bias induces an increase of the error after DR for both time and non-time redundancy cases;
- from Table 3.4b it is possible to notice that Welsch and Correntropy generally reduce the errors slightly less than the other estimators, except when bias increases. This could be due

by their redescending behaviour, making Welsch and Correntropy more sensitive to the quality of the problem initialization; the measurements adopted as initial guesses may not ensure a good quality. On the other hand, WLS reduces slightly more the errors in all scenarios except in Scenario 6. As before, this is quite unexpected.

3.1.2.2 Gross error detection

The indices AVTI and OP are reported in Appendix for all the combinations of estimators, scenarios and formulations. The main outcomes are:

- for each estimator the performance in both the cases with and without time redundancy is similar. In general, for the AVTI there is a reduction when not adopting time redundancy but also for the OP. However, for specific combinations of estimators and Scenarios this was not verified;
- for both formulations the GED performances decrease (higher AVTI and lower OP) as bias magnitude increases. This is caused by the poor reconciliation quality which impacts on the value of the residuals after DR.

Furthermore, the use of the two sets of cut-points ("low" and "high") and X84 allowed to observe that:

- the lowest AVTIs (*i.e.*, the worst performance) are achieved by applying the "high" set of cut-off points for all the scenarios and estimators. The highest AVTI values are found for the cases with bias;
- in the case of time redundancy, the highest OP (*i.e.*, the best performance) was achieved by applying X84 in Scenarios 3 and 4 and by applying the "lower" cut-off points for the two scenarios with bias.
- in the case without time redundancy, the highest OP is achieved with X84.

These observations are reasonable. A lower magnitude of the critical value for GED means a higher number of detections; it could either generate a higher number of false detections but also a higher number of right detections.

In order to provide a summary of the GED results, Table 3.5 shows the estimators with the best trade-off between AVTI and OP, considering the different Scenarios and DR approaches:

- in the case with temporal redundancy (Table 3.5a) the redescending behavior of Welsch and Correntropy provided a better GED in all the scenarios and their AVTI resulted always equal to 0;

- in the case without temporal redundancy (Table 3.5b), the monotone estimator offers good performances. In particular, Welsch and Correntropy increase their AVTI contrarily to the others estimators for which AVTI decreases with respect to the redundancy case. This can be justified by the already-mentioned issue regarding the initialization of the redescending estimators. However, Welsch and Correntropy always yield the best values.

It is clear that Welsch and Correntropy are the most promising estimators for GED in most of the scenarios. This superiority is not confirmed according to the DR results instead, except when bias increases. However, the trade-off between DR and GED has to be considered and these two estimators offer the best one. Regarding the GED methods, cut-off points can be considered more suitable than X84. In fact, they yield lower values of the AVTI index which is more relevant with respect to OP when adopting robust estimators. The estimators' robustness already copes with the presence of gross errors which could not be seen as outliers from the estimators' point of view.

Table 3.5. *Simulated steady-state case study: estimators with the best trade-off between OP and AVTI in the different Scenarios and formulations. W is for Welsch, Co for Correntropy, F for Fair, Q for Quasi Weighted Least Squares and WLS for Weighted Least Squares.*

(a)

With time redundancy			
Scenario	**AVTI c.p. "high"**	**OP c.p. "low"**	**OP X84**
1	W, Co, F	Not available	Not available
2	W, Co, F	Not available	Not available
3	W, Co, F	/	W, Co, F
4	W, Co, F	/	W, Co, F
5	W, Co		/
6	W, Co		/

(b)

Without time redundancy		
Scenario	**AVTI c.p. "high"**	**OP X84**
1	W, Co, F	Not available
2	W, Co, F	Not available
3	Q, F, WLS	
4	W, Co, F	
5	W, Co, WLS	
6	Q, WLS	

3.2 Results for the simulated dynamic case study

3.2.1 Evaluation of the data reconciliation performance

Only DR was performed for this case where Equation (1.20) was used as objective function while the differential equations (2.7), (2.8) and (2.9) were adopted as system constraints.

The non-linear dynamic constrained optimization problem was implemented and solved in the Python Optimization Modeling Objects (`Pyomo`) software. An abstract model was created considering the single time window of 50 [s]: `pyomo.DAE` modelling was used to incorporate the differential algebraic equations of the system constraints into the model. The derivatives of the differential equations were discretized according to the `pyomo.DAE` Collocation discretization scheme (Hart et al., 2017), where: number of finite elements points $nfe = 10$; number of collocation points $ncp = 5$. The non-linear solver adopted was the Interior Point OPTimizer (IPOPT) (Hart et al., 2017), and the problem was initialized using the median (for each variable) of the measurements along the entire time horizon.

3.1.2 Results

Two redescending estimators, Welsch and Correntropy, yield the highest SSE (Table 3.6a). Furthermore, their values are quite inhomogeneous respect to the other estimators. In particular, the highest SSE values are unexpectedly found in Scenario 2, where only noise affects the data. In the same scenario, TER (Table 3.6b) shows similar results, where QWLS and Fair (the monotone estimators) present the highest values. On the other hand, Welsch and Correntropy do not perform well.

As matter of example, the time-profiles of the reconciled values are shown in Figures 3.1 and 3.2, along with those of the true values and measurements. In these Figures variables $T_{e,in}$ and $\dot{V}_o$ are considered in Scenarios 3 and 6, which are the most problematic due to the high number of outliers and high bias magnitude, for Correntropy, Fair and WLS.

For Scenario 3 (Figure 3.1) the following results can be drawn:

- in the time instants when outliers (*i.e.*, the isolated red dots) are present the Fair estimator yields the best reconciliation, and the profile of the reconciled measurements (blue triangles) are smooth;
- WLS presents a quite smooth profile, but its shape does not resemble the continuous line of the true value time-profile;

- Correntropy reconciled profile is much more scattered respect to the other ones, especially in the correspondence of the outliers.

Table 3.6. *Simulated dynamic case study: mean SSE (a) and mean TER (b) over time of each estimator and scenario.*

(a)

Mean SSE					
Scenario	**Welsch**	**QWLS**	**Correntropy**	**Fair**	**WLS**
1	2.98	0.78	2.52	0.77	0.77
2	8.95	0.40	7.91	0.30	0.30
3	4.55	0.26	4.38	0.26	1.28
4	2.19	0.65	1.62	0.69	2.12
5	2.63	1.27	2.40	1.31	1.56
6	4.60	0.88	5.09	0.93	0.97

(b)

Mean TER					
Scenario	**Welsch**	**QWLS**	**Correntropy**	**Fair**	**WLS**
1	-0.98	40.62	5.85	40.89	40.58
2	-13.33	71.87	3.70	73.94	75.15
3	7.08	61.84	18.47	62.21	12.81
4	19.77	41.98	34.51	51.47	18.81
5	27.89	33.61	31.70	32.44	25.55
6	12.49	43.56	10.89	41.42	29.22

For Scenario 6 (Figure 3.2), the following results can be drawn, instead:

- Correntropy presents the same issues as in Scenario 3 when outliers are present. However, its reconciled profiles better match the true ones with respect to Fair and WLS;
- Fair and WLS reconciled values appear as biased profiles of the true values, with an increased difference between the true and reconciled profile for variable $T_{e,in}$.

There is some correspondence with what observed in the steady-state case. In presence of bias Correntropy (as well as Welsch) offers a better reconciliation quality than the other estimators. However, their performance is deteriorated by the difficulties encountered when reconciling the ethane flowrate $\dot{V}_e$. This issue also persists in the other scenarios: $\dot{V}_e$ is the variable most responsible for the lower DR quality of Welsch and Correntropy.

The reconciled value of the two inlet temperatures ($T_{e,in}$ and $T_{o,in}$) is another element of concern. As previously observed the sinusoidal profile of $T_{e,in}$ does not the match the one of the true values, as well as $T_{o,in}$. This situation happens in all scenarios and for all estimators.

These unexpected behaviours could be caused by something happening in the constraints, where the temperatures are adjusted in order to accommodate a better reconciliation of the flowrates. This situation could be explained in turn by performing the dynamic DR in a single and fixed time window, leading to some inaccurate reconciled values. Therefore, the observations drawn from this case study may need to be reviewed

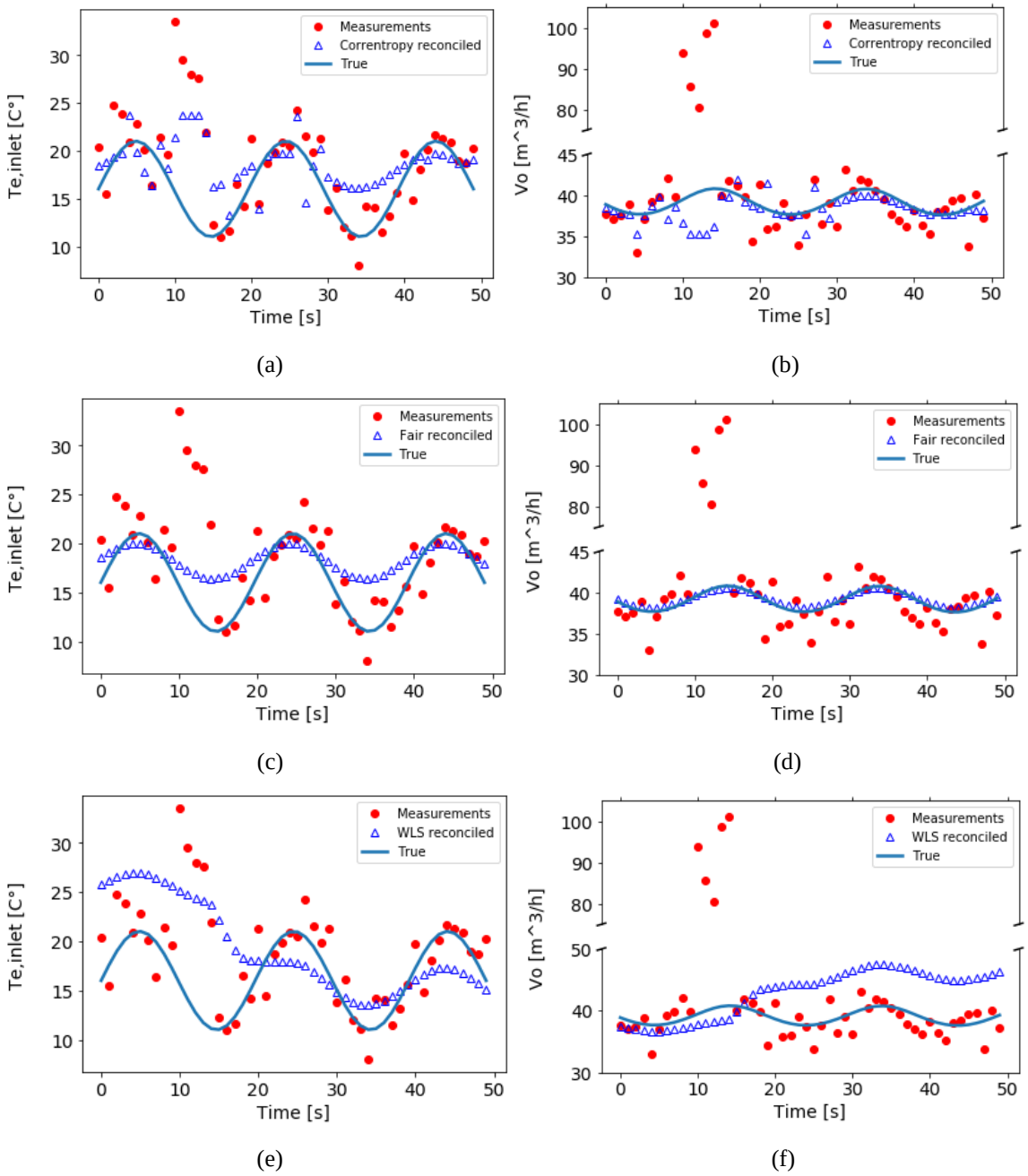


Figure 3.1. *Simulated dynamic case study: $T_{e,in}$ and $\dot{V}_o$ time profiles of true, measurements and reconciled values for Correntropy (a, b), Fair (c, d) and WLS (e, f) estimators in Scenario 3.*

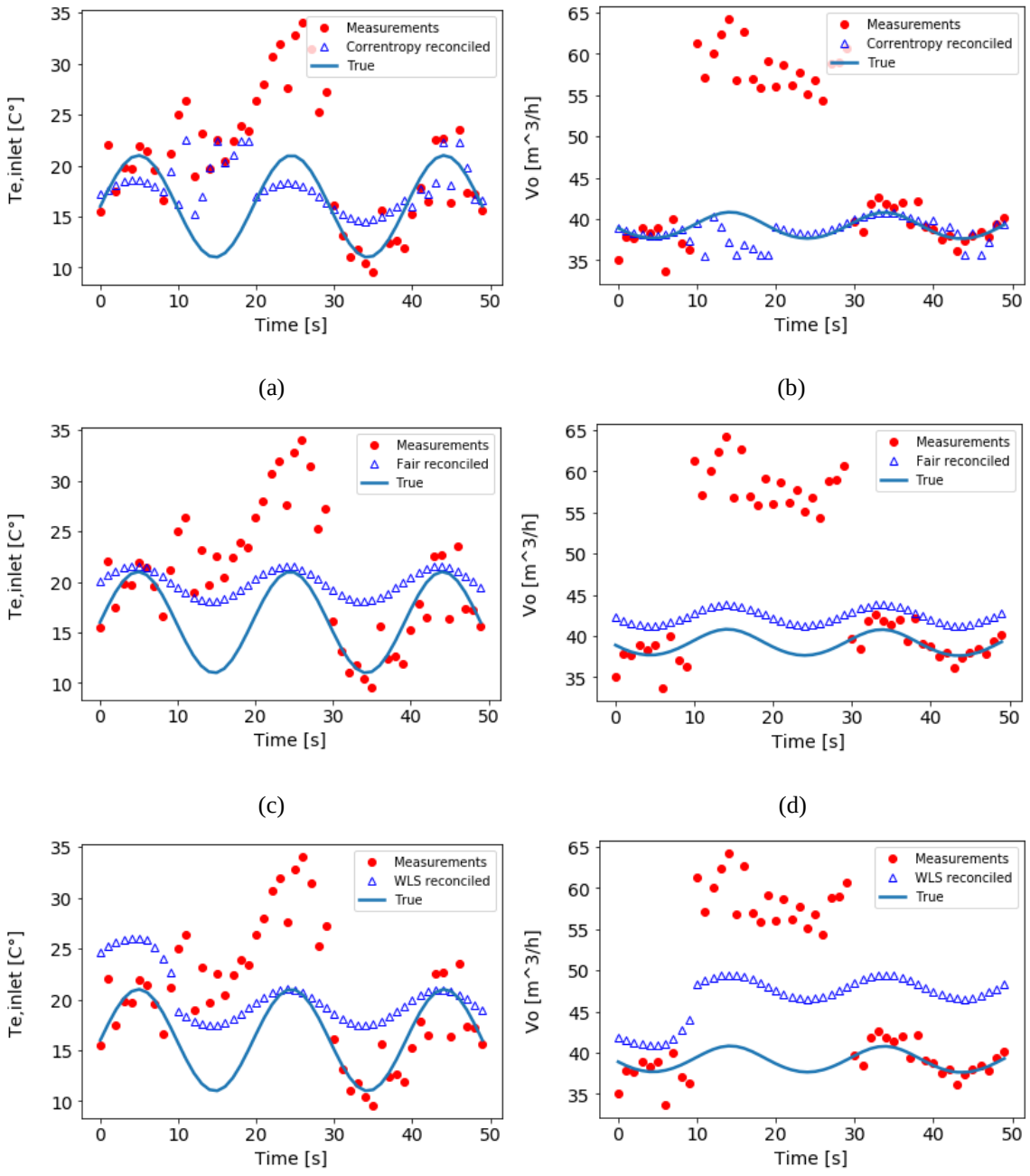


Figure 3.2. *Simulated dynamic case study: $T_{e,in}$ and $\dot{V}_o$ time profiles of true, measurements and reconciled values for Correntropy (a, b), Fair (c, d) and WLS (e, f) estimators in Scenario 6.*

CHAPTER 4

Data reconciliation and gross error detection in an industrial network of heat exchangers

In this Chapter the outcomes of the application of DR and GED to an industrial case of a heat exchangers network are shown.

4.1 Performance of data reconciliation and gross error detection in an industrial network of heat exchangers

In this Chapter only Welsch and Correntropy estimators are applied, exploiting the preliminary selection of the most promising estimators in the simulated steady-state case. DR is performed by adopting the formulation without time redundancy of Equation (1.19). This is highly beneficial because the pre-heat train inputs change rather frequently due to fouling, cleaning operations, and variability of the inlet crude oil, conditions that influence the system steady-state. Therefore, reconciling a single daily snapshot would yield final estimates closer to the steady-state conditions of the considered day.

The fact that some variables are not measured requires the resolution of the DR problem on a reduced formulation Equation (1.12); the $\mathbf{A}$ and $\mathbf{C}$ matrices of Equations (2.25) and (2.26) were divided into their unmeasured and measured submatrices, according to Equation (1.14), and the latter was set as DR constraints.

The variance percentage ratio θ was calculated in order to assess the DR quality, in absence of nominal values for SSE and TER:

$$\theta_i = 100\,\frac{\sigma_{R,i}^2 - \sigma_{msr,i}^2}{\sigma_{msr,i}^2}\,. \tag{4.1}$$

where $\sigma_{R,i}^2$ and $\sigma_{msr,i}^2$ are the variance of the reconciled values and the variance of the measurements for a certain measured variable i (where in this case $i = 1, 2, \ldots 24$),

respectively. In this way, it is possible to evaluate how much of the measurements' variability, due to random and gross errors, is reduced thanks to DR.

Before reconciling the dataset, all the measurements were scaled according to the following:

$$x_{ij}^{scaled} = \frac{x_{ij} - x_i^{min}}{x_i^{max} - x_i^{min}}. \quad (4.2)$$

x_i^{min} and x_i^{max} are respectively the minimum and the maximum value datapoints for the variable i. This data pre-treatment is necessary in order to avoid a sort of "preferential" reconciliation; the variables with the highest magnitudes could be better adjusted as they are more problematic than those with lower magnitudes (performing DR with a non-scaled dataset yielded worse results, not shown here for the sake of conciseness).

The DR of the network flowrates was performed prior to the HEN energy balances in order to have good reconciled in the energy balances reconciliation.

The network model and the DR algorithm were implemented[4] using Python. The same SLSQP method of the simulated case was used to solve the constrained optimization problem which was initialized through the measurements of the correspondent snapshot.

The maxima of the estimators' influence function and the inflection points are applied as GED method, exploiting what emerged from their comparison with X84 in the simulated steady-state case.

4.2 Results

4.2.1 Data reconciliation

Table 4.1 shows the variance ratios for the flowrates (Table 4.1a) and temperatures (Table 4.1b). The main outcomes are:

- the hot-streams flowrates FC004, FC005, FC006 and FC007 are classified as non-redundant variables after the Q-R decomposition. Therefore, they could not be adjusted through DR or in any other way. In this way, their measurements are exploited for

[4] In this Thesis robust DR and GED strategies were coded, while the network model was provided by Hexxcell.

monitoring, control, fouling prediction, etc…, as they are recorded by the sensors. The presence of possible errors can invalidate the outcomes of these operations;

- some ratios are positive meaning that the correspondent variables are bad adjusted during DR, with a consequent increase of the data variability. Fortunately, they are not too high and the greatest value found is around 6% for Welsch and around 10% for Correntropy. In both cases this value corresponds to the temperature TI011, which is the inlet hot temperature of exchanger E-05AB. This situation may occur because the correspondent measurement presents a relatively low variability (represented by the measurement's variance).

Table 4.1. *Industrial case study: variance ratio θ of flowrates (a) and temperatures (b) for Welsch and Correntropy estimators.*

(a)

Variable	*Welsch*	*Correntropy*
FC001	1.71	1.712
FC002	4.05	4.05
FC003	-10.25	-10.25
FC004	0	0
FC005	0	0
FC006	0	0
FC007	0	0

(b)

Variable	*Welsch*	*Correntropy*
TI001	-0.79	-1.31
TI002	-0.49	0.59
TI003	-19.33	-18.23
TI004	2.88	2.94
TI005	2.89	4.70
TI006	-5.97	-5.86
TI007	0.42	0.35
TI008	3.27	3.14
TI009	0.01	0.24
TI010	-7.87	-3.12
TI011	6.36	10.52
TI012	-40.57	-38.01
TI013	-0.49	1.27
TI014	3.07	5.87
TI015	0.64	0.87
TI016	-67.43	-67.50

This difference in the reconciliation quality among the variables is probably caused by the way the constraints are satisfied during the optimization. In fact, some variables are properly adjusted (*i.e.*, negative ratios are found), while others are adjusted in a less satisfactory manner

(*i.e.*, positive ratios are found) in order to fulfil the constraints. Hence, some estimates are closer to the correspondent nominal/true value of the process, due to the minimization of the error between this value and the measurement (see Equations 1.1 to 1.3), while other estimates show larger errors, meaning that they are even worse than the original measurements.

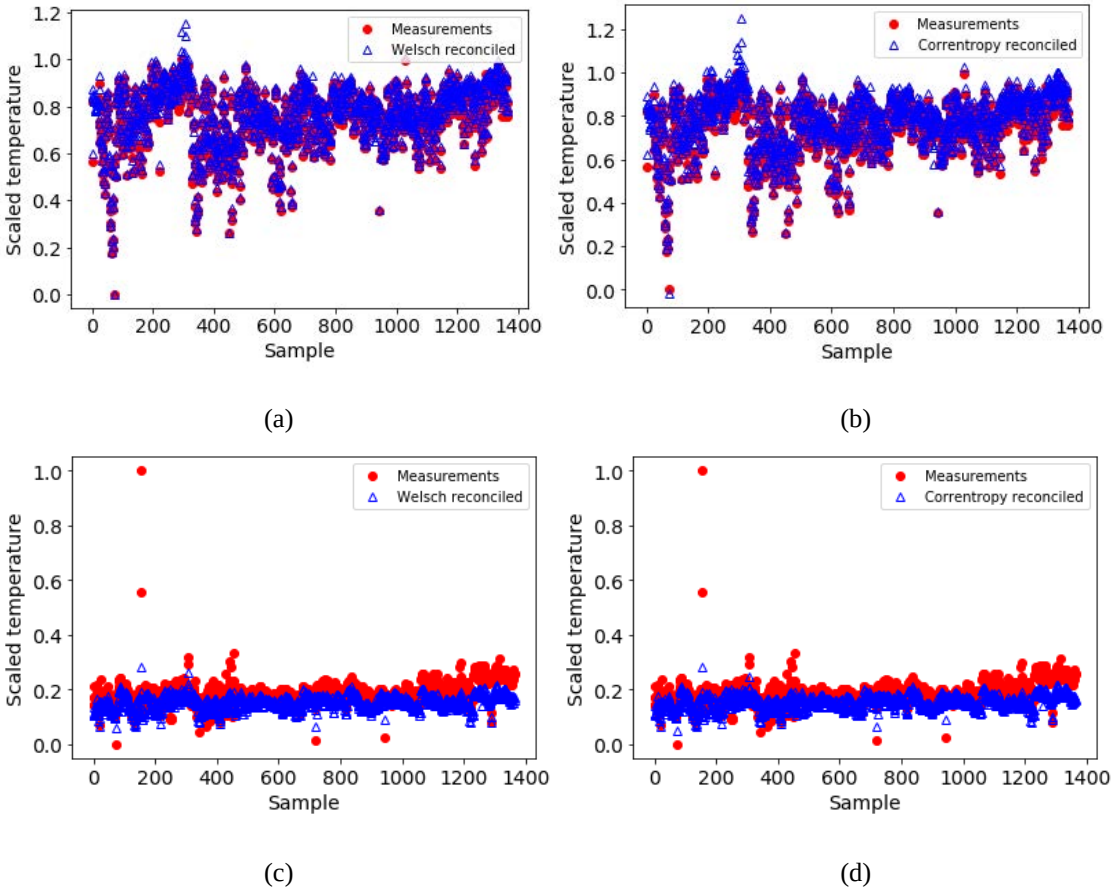


Figure 4.1. *Industrial case study: scaled measured and reconciled values of (a,b) TI011 and (c,d) TI016 for Welsch and Correntropy estimators.*

In summary, a solution is found which simultaneously satisfies the constraints and minimizes the residual. Welsch generally yields the lowest variance ratios, meaning that it offers a better reconciliation performance than Correntropy. The Welsch influence function has a smoother redescending profile respect to that of Correntropy. This difference could help to partially compensate the issues of the estimator. However, this could be only a sub-optimal solution, due to the redescending behaviour of the adopted estimators, which need a good initial guess in order to reach the optimal solution. Therefore, a better initialization strategy (*i.e.*, using

nominal values or historical reconciled value) should be adopted to improve the estimator performance.

The reconciled values profile is expected to follow the shape of the one of the measurements. The reconciled values should just smoothen the measurements profile, lowering the data-points variability and offering a better view of the true process conditions.

For example, the scaled reconciled and the measured values of TI011 and TI016 (respectively the worst and best adjusted variable) are compared in Figure 4.1 in order to verify this aspect, highlighting the results of Table 4.1 and showing the following evidences:

- in Figures 4.1a and 4.1b, the higher variability of the reconciled values of TI011 yield a more vertically stretched profile with respect to the measured ones. The original trend remains, but it cannot be confirmed if the steady-state between samples 0 and 400 effectively changes. The measurement errors and the low-quality adjustments could lead to misleading conclusions;
- in Figures 4.1c and 4.1d, the reconciled values (blue triangles) profile is shrunk and flattened with respect to the real measurements (red dots), in accordance with the low ratio for TI016, the mixer outlet temperature. In this way, it is highlighted that the steady-state of this temperature is not significantly due to possible changes in the pre-heat train inlets during time. This means that the crude flowing into the furnace has a quite constant temperature during the five years of operations.

4.2.2 Gross error detection

From the inspection of Table 4.2 the following observations are drawn:

- as previously highlighted by the simulated steady-state case, the number of detections increase when adopting lower values for the cut-off points;
- in general, Correntropy yields a slightly higher number of detections, thanks to its lower cut-off points magnitudes than those of Welsch.

The higher number of gross errors detected when using the maxima should correspond to an increase of false detections (*i.e.*, higher probability of Type I error). In order to have some confirmatory examples the inspection of the plots of the previous Figure 4.1 and of Figure 4.2 is needed. In this way it is possible to visually assess the presence of gross errors for TI006, TI011, TI012 and TI016.

Table 4.1. *Industrial case study*: *number of detected gross errors for the faulty variables.*

Variable	*Welsch max.*	*Correntropy max*	*Welsch i.p.*	*Correntropy i.p.*
FC002	0	1	0	0
TI002	2	3	0	0
TI003	5	5	4	5
TI004	2	4	0	0
TI005	3	4	0	0
TI006	4	5	0	0
TI011	0	1	0	0
TI012	9	9	6	6
TI013	7	6	0	0
TI014	0	1	0	0
TI015	3	3	0	0
TI016	14	17	2	2

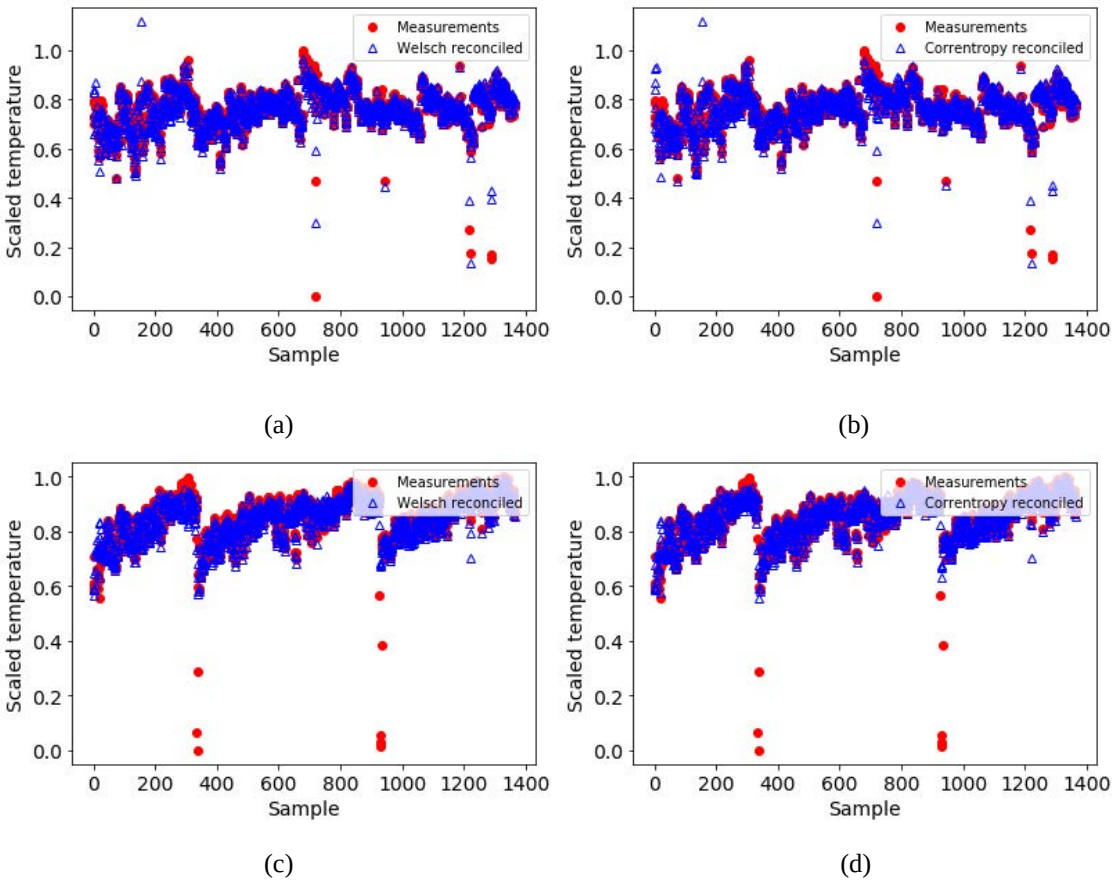


Figure 4.12. *Industrial case study: scaled measured and reconciled values of TI006 (a,b) and TI012 (c, d) for Welsch and Correntropy estimators.*

It is also advisable to report the reconciled values because through their visualization the following situations (see Equations 1.31 and 1.32) can be assessed:

- if the reconciled values are quite high respect to the correspondent measurement, they may raise the residual after DR over the cut-off points, generating a false detection. The measurement is flagged as containing a gross error when it is not. However, this situation is an issue only if the measurement has a higher magnitude than the possible nominal/true value;
- if the reconciled values are similar to the correspondent measurement (poor-quality reconciliation or slightly bad adjustments), the residual after DR is very low even though the measurement could contain a gross error, generating a missed detection.

From Figure 4.1 it can be observed that:

- for TI011, Figures 4.1a and 4.1b, the lowest red dots should be gross errors, highlighting possible process drifts. The close presence of their reconciled values disables their detection (missed detections). However, in Table 4.1 a single outlier is detected by the Correntropy maximum. It is due to the highest triangle of Figure 4.1b, generating a false detection;
- for TI016, Figures 4.1c and 4.1d, the two highest red dots are the detected process outliers by the inflection points. The most of errors found by adopting the maxima are likely to be false detections. They refer to the red dots slightly detached from the average of the measurements in both the figures. Furthermore, the three lowest red dots are probably outliers, but they are not detected due to their close adjustments.

From Figure 4.2 it can be observed that:

- for TI006, Figure 4.2a and 4.2b, the negative effect bad adjustment is highlighted. All the possible outliers have close reconciled values and the resultant detection are caused by the highest blue triangles. By adopting the inflection points these false detections are avoided, but the true gross errors are still missed;
- for TI012, Figure 4.2c and 4.2d, the red dots in the bottom part of the plots are the gross errors. In particular, the mentioned measurements were probably recorded during cleaning operations, as highlighted by the block-pattern of the measurements profile. In this way, it is possible that cleaning/maintenance operations can be seen as gross errors, when they are not. Therefore, a proper assessment is needed in order to avoid misleading detections.

From these observations it is clear how the reconciliation quality influences the GED when using cut-off points, not only leading to false detections but also to missed detections.

However, after an in-depth assessment of the obtained detections, inflection points do not generally yield false detections, providing a more robust detection; for these cut-off points Welsch and Correntropy practically offer the same results.

Conclusions

In this Thesis a strategy for robust data reconciliation and gross error detection was developed, in order to cope with the issues of measurements reliability in the heat-exchange networks. Two simulated case studies were initially considered: a heat exchanger in steady state conditions, and a heat exchanger in dynamic conditions. Then, the outcomes on the most promising data reconciliation and gross error detection methodologies were exploited in the industrial case study of a heat-exchangers network for the crude oil pre-heating in refineries.

In the first two cases a single shell-and-tube heat exchanger was considered. A preliminary assessment of the performance of six state-of-the-art robust estimators was evaluated: Simple Method, Sophisticated Method, Welsch, Quasi Weighted Least Squares, Correntropy and Fair. All these robust estimators were studied in six scenarios simulating the presence of different levels of noise in the data and the presence of different gross errors. The data reconciliation performances were evaluated through the sum of squared errors (SSE) and total error reduction (TER); the gross error detection performances were evaluated through the average number of Type I Error (AVTI) and overall power (OP). In general, considering the trade-off between data reconciliation and gross error detection, Welsch and Correntropy demonstrated to be the most promising estimators. They yielded the same SSE and TER values of the other estimators, except when bias is present where they offered by far the best results; furthermore, their AVTI and OP values were the best in most of the scenarios. The reconciliation quality was improved when exploiting temporal redundancy. However, the non-time redundancy formulation is more suited for systems where the inputs change rather frequently. In gross error detection the "cut-off points" methodology yielded a lower number of false detections.

In the dynamic case study, only data reconciliation was performed. The two monotone estimators, Quasi Weighted Least Squares and Fair, ensured the best data reconciliation performances. In fact, they always yielded the lowest SSE and to the highest TER values. However, some concerns arose regarding the quality of the results. The use of a single and fixed time window strategy for dynamic DR could be the cause of this issue. A possible improvement would be the implementation of a sliding-moving window approach (Prata et al., 2008; Prata et al., 2010).

The results obtained in the simulated steady-state case were then exploited in the real case study, where Welsch and Correntropy estimators were utilized for the reconciliation. In particular, the non-time redundancy formulation was applied in order to account for the effect of cleaning/maintenance and fouling on the network inlets. Considering the combination of DR and GED results, the Welsch estimator was the most promising one, even though the difference with Correntropy is small. In GED performed using the "cut-off points" the results were heavily influenced by the reconciliation quality, which was found to be crucial in order to have reliable gross error detection. However, the use of inflection points as "cut-off points" did not generally yield to false detections. The measurements adopted as initial guesses could be the cause for DR poor results: better candidates could improve the reconciled values quality. Furthermore, a refinement of the GED strategy could also be considered in order to be less influenced by the DR results.

A future extension of the algorithm will be oriented to include the methodology proposed by Llanos et al. (2017) to provide a systematic gross error classification for outliers, bias and drifts and take corrective actions.

Appendix

Table 2. *Simulated steady-state case study: AVTI and OP values of all estimators for DR with time redundancy: (a) Scenario 1, (b) Scenario 2 and (c) Scenario 3.*

(a)

Scenario 1			
Estimator	AVTI c.p. "low"	AVTI c.p. "high"	AVTI X84
SiM	78	78	85
SoM	78	78	85
Welsch	75	0	93
QWLS	117	59	93
Correntropy	82	0	93
Fair	70	0	93
WLS	117	59	93

(b)

Scenario 2			
Estimator	AVTI c.p. "low"	AVTI c.p. "high"	AVTI X84
SiM	78	78	85
SoM	78	78	85
Welsch	75	0	93
QWLS	117	59	93
Correntropy	82	0	93
Fair	70	0	93
WLS	117	59	93

(c)

Scenario 3						
Estimator	AVTI c.p. "low"	AVTI c.p. "high"	AVTI X84	OP c.p. "low"	OP c.p. "high"	OP X84
SiM	42	42	48	67.50	67.50	93.33
SoM	42	42	48	67.50	67.50	93.33
Welsch	40	0	52	66.67	2.50	94.17
QWLS	54	26	52	69.17	63.33	94.17
Correntropy	49	0	52	67.50	3.33	94.17
Fair	39	0	52	66.67	31.67	94.17
WLS	54	26	52	69.17	63.33	94.17

Table 2. *Simulated steady-state case study: AVTI and OP values of all estimators for DR with time redundancy: (a) Scenario 4, (b) Scenario 5 and (c) Scenario 6.*

(a)

Scenario 4						
Estimator	AVTI c.p. "low"	AVTI c.p. "high"	AVTI X84	OP c.p. "low"	OP c.p. "high"	OP X84
SiM	63	63	83	27.50	27.50	92.50
SoM	63	63	83	27.50	27.50	92.50
Welsch	62	0	84	26.67	0.00	94.17
QWLS	99	48	83	35.00	22.50	94.17
Correntropy	70	0	84	31.67	0.00	94.17
Fair	58	0	83	25.83	0.00	94.17
WLS	99	48	83	35.00	22.50	94.17

(b)

Scenario 5						
Estimator	AVTI c.p. "low"	AVTI c.p. "high"	AVTI X84	OP c.p. "low"	OP c.p. "high"	OP X84
SiM	256	256	56	7.95	7.95	5.89
SoM	257	257	56	7.95	7.95	5.89
Welsch	126	0	53	47.40	1.23	11.92
QWLS	9	6	53	17.12	13.01	12.19
Correntropy	136	0	53	48.77	1.37	11.92
Fair	122	9	53	46.85	15.34	11.92
WLS	9	6	53	17.26	12.88	11.92

(c)

Scenario 6						
Estimator	AVTI c.p. "low"	AVTI c.p. "high"	AVTI X84	OP c.p. "low"	OP c.p. "high"	OP X84
SiM	466	466	361	6.99	6.99	5.48
SoM	467	467	361	6.99	6.99	5.48
Welsch	373	0	59	83.84	19.32	50.41
QWLS	5	0	58	31.10	23.97	50.55
Correntropy	394	0	59	85.48	22.33	50.68
Fair	359	64	59	83.70	55.48	50.41
WLS	5	0	59	30.68	23.15	50.55

Table 3. *Simulated steady-state case study: AVTI and OP values of all for DR without time redundancy: (a) Scenario 1, (b) Scenario 2 and (c) Scenario 3.*

(a)

Scenario 1			
Estimator	AVTI c.p. "low"	AVTI c.p. "high"	AVTI X84
SiM	14	14	70
SoM	13	13	83
Welsch	14	2	85
QWLS	35	13	105
Correntropy	19	2	82
Fair	14	0	104
WLS	10	3	110

(b)

Scenario 2			
Estimator	AVTI c.p. "low"	AVTI c.p. "high"	AVTI X84
SiM	18	18	93
SoM	15	15	98
Welsch	15	1	105
QWLS	36	14	122
Correntropy	21	1	105
Fair	16	1	123
WLS	10	3	110

(c)

Scenario 3						
Estimator	AVTI c.p. "low"	AVTI c.p. "high"	AVTI X84	OP c.p. "low"	OP c.p. "high"	OP X84
SiM	2	2	160	23.33	23.33	47.50
SoM	2	2	160	22.50	22.50	46.67
Welsch	20	20	181	10.00	0.00	33.33
QWLS	2	2	165	25.83	20.00	47.50
Correntropy	20	20	170	10.83	0.83	33.33
Fair	2	0	160	23.33	13.33	47.50
WLS	0	0	186	22.50	17.50	47.50

Table 4. *Simulated steady-state case study: AVTI and OP values of all for DR without time redundancy: (a) Scenario 4, (b) Scenario 5 and (c) Scenario 6.*

(a)

Scenario 4						
Estimator	AVTI c.p. "low"	AVTI c.p. "high"	AVTI X84	OP c.p. "low"	OP c.p. "high"	OP X84
SiM	11	11	53	20.83	20.83	56.67
SoM	11	11	59	20.83	20.83	56.67
Welsch	19	9	48	17.50	0.00	53.33
QWLS	16	11	60	30.83	19.17	56.67
Correntropy	21	10	62	20.00	0.00	54.17
Fair	12	0	63	20.83	0.00	56.67
WLS	9	1	90	26.67	12.50	56.67

(b)

Scenario 5						
Estimator	AVTI c.p. "low"	AVTI c.p. "high"	AVTI X84	OP c.p. "low"	OP c.p. "high"	OP X84
SiM	227	227	224	0.00	0.00	47.81
SoM	180	180	160	0.00	0.00	0.00
Welsch	115	10	75	9.73	9.73	13.97
QWLS	44	32	155	0.00	0.00	26.85
Correntropy	119	11	92	11.10	10.96	15.75
Fair	226	32	127	0.00	0.00	22.88
WLS	11	2	92	0.00	0.00	10.14

(c)

Scenario 6						
Estimator	AVTI c.p. "low"	AVTI c.p. "high"	AVTI X84	OP c.p. "low"	OP c.p. "high"	OP X84
SiM	443	443	427	0.00	0.00	64.93
SoM	360	360	349	0.00	0.00	7.67
Welsch	219	93	205	27.53	27.53	0.68
QWLS	146	111	593	0.00	0.00	84.52
Correntropy	213	96	194	29.32	29.32	0.96
Fair	483	279	599	0.00	0.00	86.99
WLS	7	3	531	0.00	0.00	49.73

References

Albuquerque, J.S., Biegler, L.T., 1996. Data reconciliation and gross-error detection for dynamic systems. *AIChE Journal* 42, 2841–2856.

Arora N., Biegler L. T. (2001). Redescending estimators for Data Reconciliation and Parameter Estimation. *Comput. Chem. Eng*. 25, 1585.

Chen J. J. J. (2019). Logarithmic mean: Chen's approximation or explicit solution? *Computers & Chemical Engineering*, Volume 120, Pages 1-3.

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

Coletti F. and Macchietto S. (2011). Refinery Pre-Heat Train Network Simulation Undergoing Fouling: Assessment of Energy Efficiency and Carbon Emissions. *Heat Transfer Engineering*, 32(3–4):228–236.

Dennis J. E., Welsch R. E. (1976). Techniques for Nonlinear Least Squares and Robust Regression. *Proc. Am. Stat. Assoc*., 83−87.

Fair R. C. (1974). On the robust estimation of econometric models. *Annals of Economic and Social Measurement*, 3, 667–677.

Fuente M. J., Gutierrez G., Gomez E., Sarabia D. and de Prada C. (2015). Gross error management in data reconciliation. *IFAC-PapersOnLine*, 48-8, 623-628.

Hampel F. R., Ronchetti E. M., Rousseeuw P. J. and Stahel, W. A. (1986). *Robust statistics—the approach based on influence functions*. New York: Wiley.

Hart W. E., Laird C. D., Watson J.-P., Woodruff D. L., Hackebeil G. A., Nicholson B. L., Siirola J. D. (2017). *Springer Optimization and its Applications: Pyomo — Optimization Modeling in Python*, 2nd edition, Springer.

Johnson L. P. M., Kramer M. A. (1995). Maximum Likelihood Data Rectification: Steady-State Systems. *AIChE J*. 41, 2415.

Kongchuay P., Siemanond K. (2014). Data reconciliation with gross error detection using NLP for a hot-oil heat exchanger. *Chemical Engineering Transactions*, 39, 1087-1092 .

Llanos C. E., Sanchéz M. C, Maronna R. A. (2015). Robust Estimators for Data Reconciliation. *Ind. Eng. Chem. Res.*, 54, 5096−5105.

Llanos C. E., Sánchez M. C. and R. A. **(**2017). Classification of Systematic Measurement Errors within the Framework of Robust Data Reconciliation. *Ind. Eng. Chem. Res.* 56 (34), 9617-9628.

Loyola-Fuentes J. and Smith R. (2020). Classification and estimation of unmeasured process variables in crude oil pre-heat trains subject to fouling deposition. *Computers & Chemical Engineering*, 137, 106779.

Loyola-Fuentes J., Jobson M. and Smith R. (2019). Estimation of Fouling Model Parameters for Shell Side and Tube Side of Crude Oil Heat Exchangers Using Data Reconciliation and Parameter Estimation. *Industrial & Engineering Chemistry Research* 58 (24), 10418-10436.

Mah R. S. H. (1990). *Chemical Process Structures and Information Flows*, Butterworth-Heinemann.

Miao Y., Gang H. Su, R. and Chu J. (2009). Industrial Processes: Data Reconciliation and Gross Error Detection. *Measurement + Control*, Vol 42/7.

Narasimhan S, Jordache C. (1999). *Data reconciliation and gross error detection: an intelligent use of process data*. Gulf Professional Publishing.

Narasimhan S. and Mah R. (1987). Generalized likelihood ratio method for gross error identification. *American Institute of Chemical Engineers Journal*, 33, 1514–1521.

Nocedal J., Wright S.J. (2006). *Springer Series in Operations Research and Financial Engineering: Numerical Optimization*, 2nd edition, Springer Science+Business Media, LLC.

Ozyurt D. B., Pike R. W. (2004). Theory and practice of simultaneous data reconciliation and gross error detection for chemical process. *Computers and Chemical Engineering* 28, 381–402.

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

Prata M., D. Pinto J. C. and Lima, E. L. (2008). Comparative Analysis of Robust Estimators on Nonlinear Dynamic Data Reconciliation. *Comput. -Aided Chem. Eng.* 25 ,501 −506.

Rey W. J. J. (1983). *Introduction to Robust and Quasi-robust Statistical Methods*. Springer-Verlang, Berlin/New York.

Ripps D. L. (1965). Adjustment of experimental data. Chemical Engineering Progress. *Symposium Series*, 61(55), 8–13.

Romagnoli J. A., Sanchez M. C. (2000). *Data Processing and Reconciliation for Chemical Process Operations*. Academic Press, San Diego, CA.

Rosenberg J., Mah R. S. H. and Iordache C. (1987). Evaluation of schemes for detecting and identifying gross errors in process data. *Industrial Engineeering Chemistry Research*, 26, 555–564.

Sanchez M., Romagnoli J., Jiang Q., Bagajewicz M. (1999). Simultaneous estimation of biases and leaks in process plants. *Computers and Chemical Engineering* 23 841–857.

Serth R. and Heenan W. (1986). Gross error detection and data reconciliation in steam metering systems. *American Institute of Chemical Engineers Journal*, 32, 733–742.

Singhmaneeskulchai P., Angsutorn N. and Siemanond K. (2013). Dynamic data reconciliation in a hot-oil heat exchanger for validating energy consumption. *Chemical Engineering Transactions*, 35, 493-498.

Tamhane A. C. and R. S. H. Mah (1985). Data Reconciliation and Gross Error Detection in Chemical Process Networks. *Technometrics*, 27:4, 409-422.

Zhang Z., Shao Z., Chen X., Wang K., Qian J. (2010). Quasi-Weighted Least Squares Estimator for Data Reconciliation. *Comput. Chem. Eng.*, 34, 154 −162.