New objective function for data reconciliation in water balance from industrial processes\textsuperscript{a}

Márcio A.F. Martins\textsuperscript{*,} Carolina A. Amaro, Leonardo S. Souza, Ricardo A. Kalid, Asher Kiperstok

Rede de Tecnologias Limpas da Bahia (TECLIM), Programa de Engenharia Industrial (PEI), Universidade Federal da Bahia (UFBA), Rua Aristides Novis, 02, Federação, 40210-630 Salvador, Bahia, Brazil

A R T I C L E   I N F O

Article history:
Received 8 October 2009
Received in revised form 22 March 2010
Accepted 23 March 2010
Available online 27 March 2010

Keywords:
Data reconciliation
Uncertainty of measurement
Water balance
Quality of information
Industrial processes

A B S T R A C T

The promotion of a better rational use of water in industry requires an adequate knowledge of aqueous streams inside industrial plants. The main difficulty found in this respect is a lack of accurate measurement of such streams. In order to overcome this difficulty, a water balance tool with data reconciliation is used. This paper presents a method based on the idea that an estimated assumption can be made for any flow rate based on the best available information (quality of information). The quality of information is inversely proportional to the uncertainty of the estimated or measured flow rate and therefore there is a proportionality constant that relates these two. This paper demonstrates that the values of the proportionality constants in the real system are proximally the same for any stream. Furthermore, the proposed method has been applied to real systems with very poor measurement. This method enables the data reconciliation from water balances without data redundancy and without having to ignore flow rates with high levels of uncertainty.

\textsuperscript{a} This document is a collaborative effort.

\* Corresponding author. Tel.: +55 71 3283 9800.
E-mail addresses: marciomartins@ufba.br (M.A.F. Martins), carol.amaro@gmail.com (C.A. Amaro), leosouza@ufba.br (L.S. Souza), kalid@ufba.br (R.A. Kalid), asher@ufba.br (A. Kiperstok).

1. Introduction

The availability of the planet’s fresh water is rapidly diminishing due to the extreme increase in the consumption of water and the pollution of bodies of water. This has prompted industries, which according to World Water Vision (World Water Council, 2000) are responsible for about 20\% of the world-wide consumption of water, to adopt strategies of management of the water resources. Most industries produce some wet waste although recent trends in the developed world have been to minimize such production or recycle such waste within the production process. As regulatory issues, environmental sensitivity, and long-term water supply concerns increase, industries are considering ways to reduce their water discharges.

Over the past decades, numerous wastewater reduction and treatment technologies have been developed and practiced in the petrochemical/refinery industries (Bagajewicz, 2000; Zbontar and Glavic, 2000) and water use surveys on industrial sites have been reported in the literature (Al-Muzaini, 1998; Féres et al., 2008; Oliveira-Esquerre et al., 2009). Water management is an indispensable tool to reduce water consumption and effluent generation, as well as to reduce manufacturing costs for industries (Gaughran et al., 2007). The acquisition of water flow data is one of the great difficulties faced when constructing water balances in petrochemical sites that were designed many years ago. The combination of different data sources, including direct measurements, company records, expert estimates, scientific or technical literature and electronic databases may prove inconsistent during the data treatment (Tan et al., 2007). In the past careful use of natural resources was not focused on as it is nowadays and water was considered an inexhaustible resource. Furthermore, most water streams did not and still do not possess measurement devices. Therefore, measured flow streams are insufficient to enable the construction of water balances in most petrochemical plants (Oliveira-Esquerre et al., 2009).

This paper describes a method to solve this problem. It consists of implementing water balances associated to data reconciliation based on attributing degrees of reliability to the aqueous flow rate accordingly to empirical knowledge about the plant. It makes use of a semi-quantitative Bayesian approach to the information available (Lira and Wöger, 2006; Lira, 2007). Degrees of reliability are attributed to the data and called quality of information (Q). The main advantage of this method is that it enables the construction of water balances with any available data without initially worrying about the unknown streams. It is therefore useful for identifying water flow rates even in cases where very few streams are measured.
2. Clean technology network of Bahia programme

“Sustainable development, due to its philosophical and multidisciplinary and long-term horizons, requires a new set of visions, paradigms, policies, methodological tools and applicable procedures to be developed, tested and widely applied, if we are to make the urgently needed changes from unsustainable to sustainable societal patterns.” (Bonilla et al., 2010)

The Clean Technology Network of Bahia programme (TECLIM) was conceived in 1997 by the Department of Hydraulics and Sanitation and the Department of Chemical Engineering at the Federal University of Bahia, Brazil (TECLIM/UFBA − www.teclim.ufba.br). The main objective of TECLIM is to promote cleaner practices in industrial processes through the creation of a network linking industry, government and universities. The workgroups of TECLIM are divided into the following areas: 1) environmental management for the transfer of cleaner technologies 2) information systems 3) environmental optimization of processes 4) development of bio-indicators to assess the toxicity of industrial waste-water and 5) education and training. The details of each area as well as the structure of TECLIM are discussed in (Kiperstok, 2000). This paper is focused on the area of environmental optimization of processes. This research aims to develop and apply mathematical programming methods to reduce water consumption and industrial wastewater production. In recent years, TECLIM has carried out several research projects in partnership with industries at Camaçari Petrochemical Complex to minimize wastewater and water consumption in industries such as: Braskem (Basic Petrochemical Unit), Monsanto S.A., Deten Química S.A., Caraíba Metais S.A., Millennium S.A. These cooperative projects have enabled the development and the consolidation of a methodology to use water rationally in industries.

3. Methodology

The methodology developed in TECLIM enables the identification and management of water flows in petrochemical and chemical industries. The stages are as follows:

- Mapping the sources of water consumption and wastewater production
- Drawing up the water balance into an electronic spreadsheet
- Attribution of the QI (quality of information) for each mapped flow rate
- Detailing operational restrictions
- Applying data reconciliation to the water balance using mapped flow rates and QI.

3.1. Water balance

The water balance is used to map the aqueous streams in petrochemical and chemical units and is an useful tool for the management and use of water in the productive process (Oliveira-Esquerre et al., 2009). There are several types of streams in a water balance: clarified and distilled water, steam at varying pressures, condensed water and wastewaters.

The water balance is constructed on an electronic spreadsheet represented by diagram blocks. Each block represents a production unit, area or a piece of equipment and the lines that connect them, the aqueous input and output flows. All streams of the water balance are described with their respective flow rate and QI. Furthermore, the sum of all flow rates in each block diagram is calculated to check any existing imbalance. As the water balance is often not closed in industry, an useful tool to improve the quality data of the balance is needed. Thus, data reconciliation is the most commonly used technique to improve the data quality in any industry mass and energy balance and gross error detection (Maquin et al., 2000; Narasimhan and Jordarche, 2000; Plácido and Loureiro, 1998; Romagnoli and Sanchez, 2000).

3.2. Data reconciliation of water balance

The total considered flow rates of the input and the output streams in industrial units usually returns a figure other than zero, indicating the non-closure of the water balance. Data reconciliation is used to adjust the stream flow rates to enable the balance to close. A typical formula for aqueous streams data reconciliation problems, proposed by Crowe et al. (1983), is given by:

\[
\min \left\{ f(x, \tilde{x}) = (x - \tilde{x})^T V^{-1} (x - \tilde{x}) \right\}
\]

Subject to \( g_k(x, \tilde{x}) = 0 \)

Where: \( x \) is a vector of measured variables, \( \tilde{x} \) is a vector of reconciled value of the measured variables, \( V^{-1} \) is an inverse of the covariance matrix of measurement and \( g_k \) defines the set of model constraints (equality and inequality constraints).

This formula in mathematics is known as solving the constrained least-square optimization problem, as discussed in Bai et al. (2007). In water balance Eqs. (1) and (2) can be written without the matrix form because the measured variables are mutually independent from a statistical viewpoint, i.e. the covariance matrix \( V \) is considered to be diagonal. Then, these are given by:

\[
\min \left\{ f(q_{R}, q_{M}) = \sum_{i=1}^{N} \frac{(q_{R} - q_{M})^2}{u_{i}^2} \right\}
\]

Subject to \( \sum_{i=1}^{1} (q_{R})_{j} - \sum_{i=1}^{M} (q_{R})_{j} = 0 \) \( j = 1 \ldots J \)

Where: \( q_{R} \) = reconciled mass flow rate; \( q_{M} \) = measured mass flow rate; \( u_{i} \) = standard uncertainty of measurement associated to the measured variables; \( i \) = refers to the streams; \( N \) = involved streams; \( J \) = quantity of units involved into industrial water balance; \( L \) = input streams in the each unit; \( M \) = output streams in the each unit; \( j \) = referring to the each industrial unit. The standard uncertainty of measurement (\( u_{i} \)) is based on standard deviation from the probability density function (pdf) of the measurand (measured variable), it represents the qualitative indication to assure the quality of result of measurement (BIPM et al., 2008).

However, sometimes (for example: commercial and industrials applications) it is necessary to express the uncertainty as an interval about the result of measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand. This fraction is chosen based on required level of confidence, it represents the coverage probability that should be attributed to the measurand, its values are usually chosen to be approximately 90, 95 or 99%. Therefore, the expression of this interval to measurand is called the expanded uncertainty (\( U \)). This is multiple of standard uncertainty, therefore in terms of expanded uncertainty Eq. (1) should be written as:

\[
\min \left\{ f(q_{R}, q_{M}) = \sum_{i=1}^{N} \frac{(q_{R} - q_{M})^2}{U_i^2} \right\}
\]

Eqs. (1), (3) and/or (5) both represent the objective function of the optimization problem while Eqs. (2) and (4) both represent the
equality restrictions of the problem. These restrictions consider the mass balance of the studied control volumes. However, there is no redundancy measurement in water flow industrial plants and consequently the uncertainty of measurement is unknown. In this work a new method based on the quality of information of mapped flow rates of the water balances is used in data reconciliation problems.

3.3. Quality of information

The lack of water flow rate metering in the petrochemical industry makes the use of conventional water balances associated to classical data reconciliation unfeasible for managing water use. Because of this a new approach is needed to estimate the stream flow rates. This work considers the best available information and attributes a value to each considered stream flow rate. Moreover, according to the sources of information used, each value is either closer to or more distant from the real values. To deal with this TECLIM uses the concept of quality of information (Qi) to represent the level of uncertainty related to the stream flow rates. Unlike classical data reconciliation which requires measured flow rate, the proposed methodology does not and, therefore the flow rates obtained to construct the water balance are classified as mapped flow rates instead of measured flow rates. The relationship between Qi and the measurement uncertainty is represented by the following expression:

\[ Q_i = k_i \frac{q_{Ri}}{U_i} \]  

(6)

Where: \( i = \) refers to the streams; \( k_i = \) proportionality constant; \( q_{Ri} = \) mapped flow rate; \( U_i = \) expanded uncertainty of measurement; \( Q_i = \) quality information of the mapped flow rate. In this work mapped flow rate instead of measured flow rate will be used. Thus, rearranging and substituting Eq. (6) into Eq. (5):

\[
\min \left\{ f(q_{M}, q_{R}) = \sum_{i=1}^{N} \left[ q_{M} - q_{R} \right]^2 \frac{Q_i^2}{(k_i q_{M})^2} \right\} 
\]

Should \( k_i \) be proximally equal for all streams (as it is shown later in this paper) it could be considered as a constant in the function, so it can be removed from the sum. Eq. (7) can therefore be written as:

\[
\min \left\{ f(q_{M}, q_{R}) = \sum_{i=1}^{N} \left[ q_{M} - q_{R} \right]^2 \frac{Q_i^2}{(q_{M})^2} \right\} 
\]

Although the absolute value returned by Eq. (8) will be different from that obtained by Eq. (7), this is not meaningful for the optimization problem as the values of the variables representing the optimum point will be the same, satisfying the data reconciliation. Thus, Eq. (8) represents the new objective function in data reconciliation subject to the same restrictions of mass balance (Eq. (4)). In addition, these restrictions of balance, it considers the non-negativity restrictions (to avoid negative reconciled flow rate values) and operational restrictions (in some cases a flow rate must be higher or lower than a certain flow rate value imposed by the process).

The aqueous stream flow rates are collected by various means: measurement (if any), mass balance calculations, design documentation and information based on the opinion of plant operators.

Fig. 1. Quality of information scale.

Fig. 2. Industrial process flow diagram.
and engineers. To reconcile the water balance using these mapped flow rate values, one needs to attribute a certain degree of reliability to each of them. The concept of quality of information was used for this, i.e. each source of information is attributed a QI value, Fig. 1 shows the QI scale used.

In the proposed method, the maximum QI value (10.0) is given to the best source of information available, for example, if information is obtained from a periodically calibrated measurer (this is unlikely), data of a validated numerical simulation or the source of information obtained from reliable technical–operational report, it is classified as reliable information (RI). The minimum QI value (0.4), however, is given to flow rate values with misleading information or without registered information about the aqueous streams. The flow rate value can usually be obtained from the opinion of a plant technician or operator; for example, or even obtained from a plan without any calibrated measurer and, it is classified as misleading information (MI).

The intermediary QI value (6.0) is attributed to cases with precarious information about the aqueous streams, for example, the flow rate is obtained from a doubtful calibration measurer or design documentation and, it is classified as precarious information (PI). Therefore, this procedure enables the water balance to be constructed with any information available unlike the classical methodology which requires a set of measurements for stream flow rates.

4. Results and discussion

In this section we present the following: first the QI is presented to show how it can be used in a reconciliation data problem instead of uncertainty. After this $k_i$ (Eq. (6)) is presented as approximately constant in an industrial water balance. Finally, the results obtained with the application of this methodology in some partner industries are shown.

4.1. An example

An industrial water process system is represented by Fig. 2. The proposed data reconciliation was applied to this system. As shown in Fig. 2 for each aqueous stream a value is assumed for flow rate based on available information and QI values are assigned for these flow rates. A minimization objective function and restrictions were made considering as variables the entire stream flow rates. A routine from Brazilian free software EMSO (Environment for Modeling Simulation and Optimization) (Soares, 2006) was used to solve the optimization problem. Data available on the spreadsheet of the water balance was linked with EMSO which gave the reconciled values for the spreadsheet.

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|}
\hline
Streams & $q_M$ (kg/s) & QI & $q_R$ (kg/s) & Absolute deviation (kg/s) & Percent deviation (%) \\
\hline
F1 & 28.06 & 10 & 28.23 & 0.64 & 0.63 \\
F2 & 3.06 & 6 & 2.91 & 0.54 & 4.91 \\
F3 & 3.28 & 0.4 & 5.84 & 2.03 & 18.58 \\
F4 & 8.89 & 6 & 8.50 & 1.39 & 4.34 \\
F5 & 11.39 & 0.4 & 10.99 & 1.44 & 3.51 \\
F6 & 3.89 & 0.4 & 4.54 & 2.35 & 16.79 \\
F7 & 4.17 & 0.4 & 3.96 & 0.73 & 4.87 \\
F8 & 7.58 & 6 & 2.91 & 0.46 & 4.50 \\
F9 & 5.83 & 6 & 5.84 & 0.01 & 0.05 \\
F10 & 4.44 & 6 & 4.54 & 0.35 & 2.19 \\
F11 & 3.89 & 6 & 3.96 & 0.27 & 1.93 \\
F12 & 13.00 & 10 & 14.95 & 1.85 & 13.33 \\
F13 & 13.33 & 10 & 12.28 & 0.19 & 0.40 \\
\hline
\end{tabular}
\caption{Results of data reconciliation.}
\end{table}

Using Eq. (8) for this system the values found for the stream flow rates after reconciliation are shown in Table 1. Their absolute and percent deviations were calculated using Eq. (9) and (10) respectively.

\begin{table}[h]
\centering
\begin{tabular}{|l|l|l|l|l|l|l|}
\hline
Node & Mapped balance (kg/s) & Reconciled balance (kg/s) \\
\hline
1 & 0.56 & 0.00 \\
2 & 0.28 & 0.00 \\
3 & 0.56 & 0.00 \\
4 & 0.83 & 0.00 \\
5 & 0.28 & 0.00 \\
6 & 0.56 & 0.00 \\
7 & 0.28 & 0.00 \\
8 & 0.28 & 0.00 \\
\hline
\end{tabular}
\caption{Results of data reconciliation in each node.}
\end{table}

As expected (Table 1) the larger corrections were for flow rates with lower QI while the flow rate with a QI equal to 10 underwent very few corrections. This occurs in classical data reconciliation, i.e. larger and lower corrections will refer to flow rates with higher or lower levels of uncertainty respectively.

Table 2 shows the balance in each node before and after reconciliation. According to the values obtained after reconciliation, the closures of the balance at each node of the flow diagram are shown. Therefore, this proposed methodology generates results similar to the traditional water balance method. Moreover, TECLM’s method has further advantage: there is no need to estimate the measurement uncertainty, instead it makes a suitable choice for the source of information related to the mapped flow rate.

4.2. Calculation of the $k_i$ constant

There is an assumption in Eq. (8) that values of $k_i$ for each mapped flow rate could be considered approximately equal. To show this, a real industry system with three QI levels of instrumentation was studied. According to classic reconciliation procedures with measured flow rate data it is possible to estimate the measurement uncertainty (standard and expanded uncertainty) of the values involved using the ISO—GUM standard method (BIPM et al., 2008).

In industrial plants there are several types and qualities of instruments used to measure fluids. For this work six streams were classified in pairs as reliable information (RI), precarious information (PI) or misleading information (MI). The $QI$ attributed for each

\begin{table}[h]
\centering
\begin{tabular}{|l|l|l|l|l|l|l|}
\hline
Streams & A & B & C & D & E & F \\
\hline
Measure (MI), kg/s & 1.55 & 22.90 & 1793.37 & 2633.25 & 8.57 & 64.57 \\
Combined uncertainty & 0.07 & 1.12 & 4.98 & 7.32 & 0.01 & 0.10 \\
$(\alpha_i, \lambda_i)$, kg/s & 0.11 & 1.85 & 8.20 & 12.04 & 0.02 & 0.17 \\
Expanded uncertainty (UI), kg/s & 25.98 & 29.03 & 1.65 & 1.65 & 0.95 & 0.95 \\
Relative Uncertainty $U_i$ (%) & 0.4 & 0.4 & 6.0 & 6.0 & 10.0 & 10.0 \\
$Q_i$ (dimensionless) from mean $k_i$ & 0.094 & 0.104 & 0.099 & 0.099 & 0.095 & 0.095 \\
Relative deviation of $k_i$ & 4.1 & 6.1 & 1.0 & 1.0 & 3.1 & 3.1 \\
\hline
\end{tabular}
\caption{Comparison among $k_i$ for each QI of an industrial process.}
\end{table}
type of instrument were 10.0, 6.0 and 0.4 respectively based on the proposed QI scale. For each type of instrument the following range of ratio between uncertainty and measured flow rate (relative uncertainty $U_r$) was considered: $QI = 10.0$ the $U_r$ must be within the range 0.2%–1.0%; $QI = 6.0$ the $U_r$ must be within the range 1.0%–8.0% and $QI = 0.4$ the $U_r$ must be within the range 8.0%–30.0%.

Thus, with a knowledge of both the estimated uncertainty and the previously attributed $QI$ for each of the streams, it was possible to calculate the values of the constants $k_i$ defined by Eq. (6). These are shown in Table 3. It can be observed that the values obtained for $k_i$ converge to a value close to $k$ equal to 0.098. Considering that the mean value of $k_i$ is 0.098 (this value was obtained by the arithmetic mean of each calculated value) the absolute and relative deviation from this mean of all $k_i$ were calculated. Thus, it can be assumed that $k_i$ is approximately constant and that Eq. (8) is valid in an industrial water balance.

In this work values of $QI$ between 0.4 and 10 were attributed to the mapped flow rates. The $QI$ values were validated using the results obtained from data reconciliation which were exhaustively discussed with the technical staff of the plant. The proposed methodology was then validated by the fact that the $k_i$ values found were very close to each other.

### 4.3. Real case studies

The application of this methodology provides information about the main aqueous stream flow rates in industries sites. Furthermore, it obtains some relevant information about aqueous profile inside the plants. Some specific results obtained in some partner industries at the Camaçari Petrochemical Complex in Bahia are:

- Identification of the largest consumers of water and largest generators of effluent inside battery limits.
- Identification of several leaks in the areas which the plant staff had no knowledge of.
- Identification of the aqueous profile through environmental indicators.

Thus, gaining knowledge of these aqueous streams is a fundamental stage in achieving any reduction in water consumption and the minimization of wastewater generation in industrial processes. Furthermore, this methodology provides basic support for the identification of reuse and recycling opportunities closer to the processes. An environmental practice that lead to a more comprehensive and quantitative characterization of environmental

![Fig. 3. Profile of effluent generated and petrochemical production of Braskem/UNIB.](image)

![Fig. 4. Profile of water consumption at Lyondell.](image)
performance in plant sites are the indicators of water consumption and effluent generation (Perotto et al., 2008). The proposed methodology was responsible for building these environmental indicators inside plants at the Camaçari Complex.

A real example is presented in Fig. 3 which shows the evolution of the environmental performance at the Braskem/UNIB. There has been a reduction in the effluent generated of approximately 250 m³/h since the start of the cooperative project with Braskem. The effluent flow rate was normalized for reasons of company confidentiality.

Another real example was carried out during the cooperative project at Lyondell’s site. Before the project, Lyondell had adopted policies to reduce water withdrawal and water consumption stabilized at a certain level. With the cooperative project ongoing, water consumption has decreased by a further 11% since beginning of the project. These results are presented in Fig. 4. The data was also normalized here.

Therefore, an understanding of water flows throughout the development of the water balance associated with data reconciliation, together with other initiatives and different actions on the part of the companies have contributed to reducing water consumption and minimization the total wastewater generation. Furthermore, the obtained results of this proposed methodology enables to apply suitable techniques of mass integration for the minimization of generated effluent such as the cases presented in (Pessoa et al., 2007; Wang and Smith, 1994). The techniques of mass integration applied in industrial sites not only contribute to a reduction in effluent and water consumption but they also provide potential economic gains, according to Smith and Petela (1992). Traditional approaches to water minimization such as changing washing operations, when complemented with water pinch analysis methodology have been shown that it is possible to achieve 30–60% fresh water savings in industrial applications.

5. Conclusion and future works

Classic reconciliation data techniques cannot be used in water balance because they do not have a redundancy measurement and uncertainties cannot be quantified and/or the cost of estimating uncertainty is very high. In this paper we have shown that QI can represent a certainty measure. The techniques described in this paper enable the determination of the aqueous stream flow rates inside an industrial plant with very poor measurement. Measurement, exact or approximate, is important for the minimization of water consumption and wastewater generation.

This method is being tested and improved in other case studies. Discussions with plant staff are being carried out to improve the ways in which QI values are attributed to the mapped aqueous streams. In fact, the possibility of estimating QI’s using Fuzzy Logic is also being developed. This method enables the handling of imprecision in available information and provides a basis for a qualitative approach to the choice of the appropriate QI.

Acknowledgments

The authors gratefully acknowledge the financial support given by the FINPE (State Agency for Innovation), FAPESB (Foundation for Support for Research of Bahia) and CNPq (Brazilian Agency for Scientific and Technological Development); the Companies at Camaçari Complex Braskem/UNIB (the large company at the largest industrial complex in Latin America located in Camaçari, Bahia, Brazil) and Lyondell for their technical and financial support.

Nomenclature

- \( x \): vector of measured variables (kg/s)
- \( \hat{x} \): vector of reconciled variables (kg/s)
- \( V^{-1} \): inverse matrix of covariance of measured variables (kg/s²)
- \( U_c \): combined standard uncertainty for a measured mass flow rate (kg/s)
- \( U \): expanded uncertainty for a measured mass flow rate (kg/s)
- \( U_r \): relative uncertainty for a measured mass flow rate (kg/s)
- \( q_{m} \): mapped mass flow rate (kg/s)
- \( q_e \): reconciled mass flow rate (kg/s)
- \( Q_I \): quality of information assigned to each mapped mass flow rate (dimensionless)
- \( k \): proportionality constant which relate \( Q_I \) and \( U_r \) (dimensionless)

Greek letters

- \( \Delta \): absolute deviation in problem data reconciliation (kg/s)
- \( \Delta P \): percent deviation in problem data reconciliation (%)

References

Lira, I., Wöger, W., 2006. Comparison between the conventional and Bayesian approaches to evaluate measurement uncertainty. Metrologia 43, 249–259.