



A MINLP FORMULATION TO OPTIMIZE SENSOR ALLOCATION USING RECONCILED DATA IN SYSTEMS WITH FEWER MEASUREMENTS.

*Marcos Narciso*¹, *Edson Valle*², *Elias Braga*³, *Asher Kiperstok*⁴, *Ricardo Kalid*⁵

¹ Universidade Federal da Bahia, Programa de Pós Graduação em Engenharia Industrial, Salvador, Brazil, marcosnarciso@gmail.com

² Universidade Federal do Rio Grande do Sul, DEQUI – Departamento de Engenharia Química, Porto Alegre, Brazil, edsoncv@gmail.com

³ Petróleo Brasileiro S/A - Fábrica de Fertilizantes Nitrogenados da Bahia, Camaçari, Brazil, eliasbraga@petrobras.com.br

⁴ Universidade Federal da Bahia, Programa de Pós Graduação em Engenharia Industrial, Salvador, Brazil, asher@ufba.br

⁵ Universidade Federal da Bahia, Programa de Pós Graduação em Engenharia Industrial, Salvador, Brazil, kalid@ufba.br

Abstract: The positioning of new mass flow sensors in an industrial plant is a challenge problem, especially in water networks where measurements are rarely found. The purpose of this paper is to study the application of a methodology to obtain the optimal location for a new sensor set in water networks. To perform such task, the results of a data reconciliation using the concept of Quality of Information, *QI*, is used. First, the mapping of the network topology and flows was performed to feed a data reconciliation with *QI*. Second, with the reconciled data, some runs in objective function, with some variation in optimization parameters, were carried out to obtain the optimal positioning of the new sensor set. The optimization was performed respecting all the mass and economic constraints of the measurement system, maximizing the objective function proposed. The proposed method was applied to a real industrial unit: Fafen-BA, a Brazilian nitrogen fertilizer plant. The results of the methodology application show that it is appropriated to choose optimal location of new sensors in an industrial plant.

Keywords: Sensor network design, MINLP Optimization, Data reconciliation, Reliability, Quality of information.

1. INTRODUCTION

The increasing use of data acquisition in process plants led to the need for optimization of their performance for quality improvement and process operation. In industrial water networks, mass or volume flow sensors are rarely found because they were considered low-profit systems. However, it is important for industries to know the amount of water consumed or wasted. In this context, data reconciliation offers a good tool to decrease process uncertainty and to estimate unmeasured variables [1].

Measurement redundancy is necessary to obtain good results in data reconciliation. However, in industrial scenarios, with fewer or none measurements, it is impossible to make classical reconciliation. An alternative approach to perform data reconciliation in these systems is to apply a proposed methodology drawn up by Fontana et al. [2] and Martins et al. [3]. This uses process information from alternative sources, such as mass and energy process

balances, information provided from experienced process engineers, direct measurements and project data to solve the problem of data reconciliation.

A number relative to the data quality is attributed to each piece of data collected and it is inversely proportional to uncertainty of measurement. This number is called Quality of Information (*QI*) and a table with predetermined values for it has to be drawn up prior to running the reconciliation. A higher *QI* number is attributed to high quality information which contemplates only minor uncertainty.

This solves the problem of low, or absent, measured data to make the reconciliation. But the decreasing of uncertainty can only be consolidated by installing new sensors in the plant. Some systematic tools to perform a cost-benefit analysis of the allocation, or also an upgrade, of a new sensor have been investigated by researchers.

Two strategies were proposed by Bagajewicz and Sánchez [4]: in the first, the objective function based on minimum cost of allocation, while the other, a maximum precision design model. These were the earliest models to provide a sensor allocation using binary variables in the objective function. In this paper, the latter is used to find the optimal position for a new sensor using data provided by reconciliation with *QI*. The objective function was modified to incorporate this information. Furthermore flow data is needed to correct the unit of measurement of the system. *QI* is dimensionless.

This paper presents a new approach to obtain the optimal allocation for new sensors set with maximum precision with minimal costs in systems with fewer measured variables. In the following sections, MINLP models for sensor allocation and a methodology for data reconciliation in systems without redundancy are considered. The models presented by Bagajewicz and Sánchez [4] are reviewed. Finally, two models are proposed that minimizes the total uncertainty quantities of upgrading instrumentation, using *QI* data and the cost of sensors. One illustrative example of a real process without measurements is solved in both models and its results are discussed.

2. PROBLEM STATEMENT

The optimal cost sensor network with precision constraints is obtained solving the following optimization problem [4]

$$\left. \begin{array}{l} \min_{q_i} \sum_{i \in M_w} q_i \cdot c_i \\ \text{s.t.} \\ \hat{\sigma}_j(q) \leq \sigma_j^* \quad \forall j \in M_p \\ q_i = \{0,1\} \quad \forall i \in M_w \end{array} \right\} \quad (1)$$

where q_i is the vector of binary variables that indicates if a sensor is allocated in position i , c_i is the cost of such a sensor, M_w is the search space, M_p is the set of reference parameters to objective function, $\hat{\sigma}$ is the estimated uncertainty from data reconciliation and σ^* is a threshold value to uncertainty.

The maximum precision sensor network with cost constraints is obtained solving the following optimization problem [4]

$$\left. \begin{array}{l} \min_{q_i} \sum_{j \in M_p} a_j \cdot \sigma_j^2(q) \\ \text{s.t.} \\ \sum_{i \in M_w} c_i \cdot q_i \leq c_T \\ q_i = \{0,1\} \quad \forall i \in M_w \end{array} \right\} \quad (2)$$

where a_j is a sensibility factor for variable $j \in M_p$ and c_T is the available investment. A special theorem provided by Tuy [5] can be used to proof that the problem presented in Eq. (2) is the dual of optimal cost problem presented in Eq. (1).

Furthermore, any of these approaches can be used to solve this problem, but it is limited to linear systems. For non-linear systems, the uncertainty can be estimated by first linearizing of model on the neighborhood of the optimal point of data reconciliation.

The linearized model can be approximated by [4]:

$$\mathbf{J} \cdot \mathbf{z} \approx \mathbf{d} \quad (3)$$

where \mathbf{J} is the Jacobian matrix of $f(z)$, the mass or energy balances in plant, in the optimal solution z_0 (the expected operational point) and \mathbf{d} is the corresponding vector of constants. In turn, the uncertainty can be obtained by [1]:

$$\hat{\mathbf{S}} = (\mathbf{I} - \mathbf{J}^T \times (\mathbf{J} \times \mathbf{S} \times \mathbf{J}^T)^{-1} \times \mathbf{J})^T \times \mathbf{S} \times (\mathbf{I} - \mathbf{J}^T \times (\mathbf{J} \times \mathbf{S} \times \mathbf{J}^T)^{-1} \times \mathbf{J}) \quad (4)$$

where \mathbf{S} is the matrix of uncertainty before data reconciliation. Therefore, $\sigma_i = \hat{\mathbf{S}}_{ii}$.

In systems with a fewer measurements, the process to obtain its uncertainty can be very difficult. Fontana et al. [2] proposed and Martins et al. [3] validated a new approach to estimate uncertainty for those systems by introducing the concept of Quality of Information (QI). The QI is a

Bayesian data [6] and can be obtained from many sources, such as mass or energy balances, local measurements, and information provided by experienced engineers on plant. A score is attributed for each collected piece of information. Thus, a higher QI represents a lower uncertainty. This formulation can be represented by [1, 2]:

$$QI_i = \kappa_i \cdot \frac{z_i}{\sigma_i} \quad (5)$$

where κ_i is the proportionally constant. This modification enables the estimation of the reconciled variables without redundancy of measurements. The matrix \mathbf{S} in Eq. (4) is replaced by \mathbf{S}_{QI} calculated by the relation given in Eq. (5).

2.1. A Quality of Information MINLP model

Whereas there is no preexisting redundancy in measurements, neither of software (obtained by model computation) either of hardware (obtained by sensor measurement), and just one instrument will be installed at each position, we started to develop the model of optimal sensor allocation using QI data. A maximum reliability model will be used as the base for this kind of problem because it is easier to understand than the optimal cost model.

Introducing a change of variables proposed in Eq. (5), Eq. (2) can be rewritten as follows

$$\left. \begin{array}{l} \min_{q_i} \sum_{i \in M_w} a_i \cdot \left(\kappa_i \cdot \frac{z_i}{QI_i(q)} \right)^2 \\ \text{s.t.} \\ \sum_{i \in M_w} c_i \cdot q_i \leq c_T \\ q_i = \{0,1\} \quad \forall i \in M_w \end{array} \right\} \quad (6)$$

where z_i is the measurement variable and $QI_i(q)$ is a function of installed sensor, as well as the uncertainty function in maximum precision model by Eq. (2), a_i is the same sensibility factor, a_j , of Eq. (5) and it can be defined by the engineering.

Note that all variables are included in the objective function, rather than only process parameters. This is done due to the approach, which targets an increasing of the QI on measurement system. This QI , also known as Global QI , can be calculated by [2]

$$QI_G = \frac{\sum_i a_i \cdot z_i \cdot QI_i}{\sum_i a_i \cdot z_i} \quad (7)$$

With these modifications nonlinearity is introduced in the problem which becomes a Mixed-Integer Nonlinear Programming (MINLP) model.

2.2. A global Quality of Information MINLP model

The Global QI is an easy way to observe how to uncertainty decreases with the increase of available cost of

sensors installation [1]. With the concept introduced by Eq. (7), a new MINLP model can be written as follows

$$\left. \begin{array}{l} \max_{q_i} \quad QI_G(q) \\ s.t. \quad \sum_{i \in M_w} c_i \cdot q_i \leq c_T \\ \quad \quad q_i = \{0,1\} \quad \forall i \in M_w \end{array} \right\} \quad (8)$$

where $QI_G(q)$ is the Global QI with reconciled flows calculated by Eq. (7) and it is a function of installed sensor.

In the model represented by Eq. (8), the objective is to maximize the Global QI . This can be explained by Eq. (5), because QI is inversely proportional to uncertainty.

2.3. Numerical Aspects

Due to the formulation of these models, unmeasured variables with large variance, i.e., with $QI_i(q)$ values close or equal to zero, can cause a numerical error due to the division by zero in Eq. (4) and Eq. (5). Therefore, it is recommended to be careful in choosing the QI value aiming to prevent this event.

Also, values of z_i close or equal to zero makes the matrix S calculated by Eq. (4) singular, affecting the resolution of models. To prevent this, it is recommended to use a reference value for zero, usually a little larger than the machine or software limit of accuracy. The use of reconciled data for z_i is also a good practice because the resolution of the MINLP model will always start from an optimal point of the process, respecting its constraints.

Models were implemented in the software EMSO (acronym for Environment for Modeling Simulation and Optimization) [7] in its beta version 0.10.6, and solved using an interface with BONMIN solver [8].

3. EXAMPLE – UNMEASURED WATER NETWORK IN A NITRIC ACID PLANT

Consider the process network of Fig. 1 which represents a simplified water flow network in a nitric acid production process in Fafen-BA (Operating Unit Nitrogen Fertilizer Plant in Bahia) located in Camaçari Petrochemical Complex - Brazil. It contains twelve nodes and 26 streams, which 21 are unmeasured and five of them cannot be measured (in cases where the source is rainwater, evaporation in process or water generated/consumed in chemical reactions).

For solving assumptions, all unmeasured streams are candidates for new sensor allocation and no hardware redundancy is allowed. It is also assumed that all candidates' sensors have a QI value equal to 20 and have cost directly proportional to respective stream flow. Flows that cannot be measured are indicated in Table 1 by the null cost of sensor and are indicated in Figure 1 by blue arrows. The reconciled flow rates, the reconciled QI associated and the costs of all available sensors are shown in Table 1.

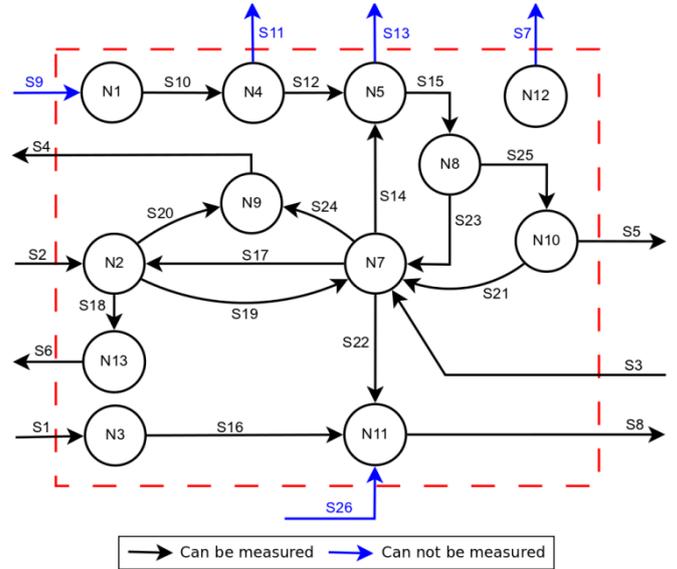


Fig. 1. Process network of Example

Table 1. Data used in Example.

| Stream | z_i (t h ⁻¹) | QI_i | c_i (US\$) |
|-----------------|----------------------------|---------|--------------|
| S ₁ | 7.14 | 2.9027 | 1000.00 |
| S ₂ | 1.91 | 6.2454 | 500.00 |
| S ₃ | 1.01 | 6.9917 | 200.00 |
| S ₄ | 0.11 | 4.3411 | 100.00 |
| S ₅ | 0.34 | 12.6442 | 100.00 |
| S ₆ | 0.68 | 3.7645 | 300.00 |
| S ₇ | 0.81 | 4.8523 | -- |
| S ₈ | 9.74 | 3.9607 | 800.00 |
| S ₉ | 1.60 | 9.6945 | -- |
| S ₁₀ | 1.60 | 9.6945 | 300.00 |
| S ₁₁ | 0.42 | 4.3747 | -- |
| S ₁₂ | 1.18 | 7.3258 | 300.00 |
| S ₁₃ | 0.52 | 6.6157 | -- |
| S ₁₄ | 1.45 | 7.7179 | 200.00 |
| S ₁₅ | 2.11 | 12.8883 | 300.00 |
| S ₁₆ | 7.14 | 2.9027 | 1000.00 |
| S ₁₇ | 6.79 | 5.1677 | 800.00 |
| S ₁₈ | 0.68 | 3.7645 | 100.00 |
| S ₁₉ | 7.96 | 6.0633 | 800.00 |
| S ₂₀ | 0.06 | 1.6941 | 100.00 |
| S ₂₁ | 0.06 | 8.1242 | 100.00 |
| S ₂₂ | 2.44 | 10.8030 | 200.00 |
| S ₂₃ | 1.83 | 11.1516 | 300.00 |
| S ₂₄ | 0.05 | 1.8748 | 100.00 |
| S ₂₅ | 0.28 | 10.4493 | 100.00 |
| S ₂₆ | 0.16 | 1.8810 | -- |

Optimizations were performed for different values of costs. For each case there were 2²¹ nodes to be explored by EMSO and BONMIN solver and execution time on an Intel Core 2 Duo PC, 2.3 GHz, 2 GB RAM ranged between 4 s (for the simplest case) and 40 s (for the most complex case). The results for each case are shown in Table 2, for the QI MINLP model, and Table 3 for the Global QI MINLP model. In this example, nonlinear models have 26 continuous variables, just 21 was used in optimization.

Considering the QI MINLP model calculated by Eq. (6) which results are presented in Table 2 and on Fig. 2, it is

possible to notice a slight decrease on the final value of sum of uncertainty for a total investment of sensor installation after US\$ 4 000.00. Similarly, there is a slight increase in Global QI beginning in US\$ 4 000.00. This fact represents a limit on the efforts in increasing of investment in new sensors. This fact is also indicated in Fig. 3 and in Table 3. In this case it is recommended to use the results for total investment cost until US\$ 5 000.00.

Table 2. Results of Example for QI MINLP Model.

| c_T (US\$) | $\sum q_i c_i$ (US\$) | $\sum u_i^2$ ($t h^{-1}$) ² | QI_G | Sensors |
|-----------------|--------------------------|---|---------|---|
| 0.00 | 0.00 | 0.2205 | 5.6875 | – |
| 1500.00 | 1500.00 | 0.1603 | 8.6733 | $S_1, S_{14}, S_{18}, S_{22}$ |
| 2000.00 | 2000.00 | 0.1021 | 9.8888 | S_1, S_{16} |
| 2800.00 | 2800.00 | 0.0440 | 12.5780 | S_1, S_8, S_{16} |
| 3550.00 | 3500.00 | 0.0428 | 13.4173 | $S_1, S_2, S_8, S_{16}, S_{22}$ |
| 4000.00 | 4000.00 | 0.0272 | 14.9092 | $S_1, S_8, S_{16}, S_{17}, S_{18}, S_{20}, S_{22}$ |
| 5000.00 | 5000.00 | 0.0111 | 16.8654 | $S_1, S_2, S_8, S_{16}, S_{17}, S_{18}, S_{19}$ |
| 6050.00 | 6000.00 | 0.0098 | 18.0075 | $S_1, S_2, S_6, S_8, S_{12}, S_{14}, S_{16}, S_{17}, S_{18}, S_{19}, S_{22}$ |
| 8000.00 | 7700.00 | 0.0091 | 19.2203 | $S_1, S_2, S_3, S_4, S_5, S_6, S_8, S_{10}, S_{12}, S_{14}, S_{15}, S_{16}, S_{17}, S_{18}, S_{19}, S_{20}, S_{21}, S_{22}, S_{23}, S_{24}, S_{25}$ |

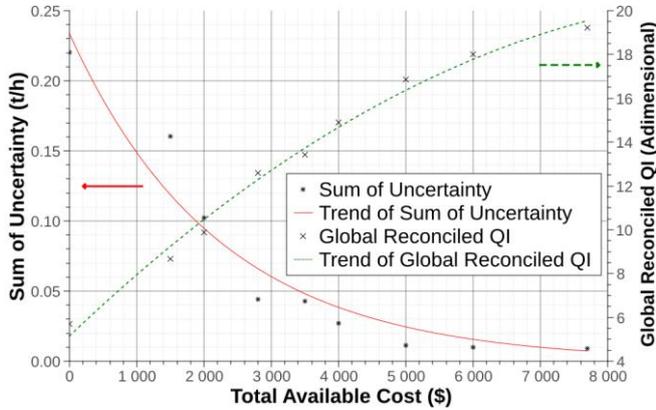


Fig. 2. Results of Example for QI MINLP Model

4. CONCLUSION

The main contributions of this work are the MINLP formulations for the sensor allocation task even when uncertainty data in measurement system is unavailable. These formulations were validated with an industrial case study with 21 unmeasured variables which was solved with satisfactory results. Future works could apply the proposed methodology could extend the application to solve other industrial systems such as heat exchangers networks systems where mass and temperature flows (energy balance) are present.

Table 3. Results of Example for Global QI MINLP Model.

| c_T (US\$) | $\sum q_i c_i$ (US\$) | $\sum u_i^2$ ($t h^{-1}$) ² | QI_G | Sensors |
|-----------------|--------------------------|---|---------|---|
| 0.00 | 0.00 | 0.2205 | 5.6875 | – |
| 1500.00 | 1500.00 | 0.1612 | 9.4884 | $S_3, S_8, S_{14}, S_{18}, S_{22}$ |
| 2000.00 | 2000.00 | 0.1460 | 10.9809 | $S_8, S_{14}, S_{19}, S_{22}$ |
| 2800.00 | 2800.00 | 0.0872 | 12.7741 | S_1, S_8, S_{19}, S_{22} |
| 3550.00 | 3500.00 | 0.0711 | 14.3124 | $S_8, S_{16}, S_{17}, S_{18}, S_{19}$ |
| 4000.00 | 4000.00 | 0.0277 | 15.1822 | $S_1, S_8, S_{14}, S_{16}, S_{19}, S_{22}$ |
| 5000.00 | 5000.00 | 0.0112 | 17.1537 | $S_1, S_8, S_{14}, S_{16}, S_{17}, S_{18}, S_{19}, S_{22}, S_{25}$ |
| 6050.00 | 6000.00 | 0.0105 | 18.1547 | $S_1, S_3, S_8, S_{10}, S_{14}, S_{15}, S_{16}, S_{17}, S_{18}, S_{19}, S_{22}, S_{23}$ |
| 8000.00 | 7700.00 | 0.0091 | 19.2203 | $S_1, S_2, S_3, S_4, S_5, S_6, S_8, S_{10}, S_{12}, S_{14}, S_{15}, S_{16}, S_{17}, S_{18}, S_{19}, S_{20}, S_{21}, S_{22}, S_{23}, S_{24}, S_{25}$ |

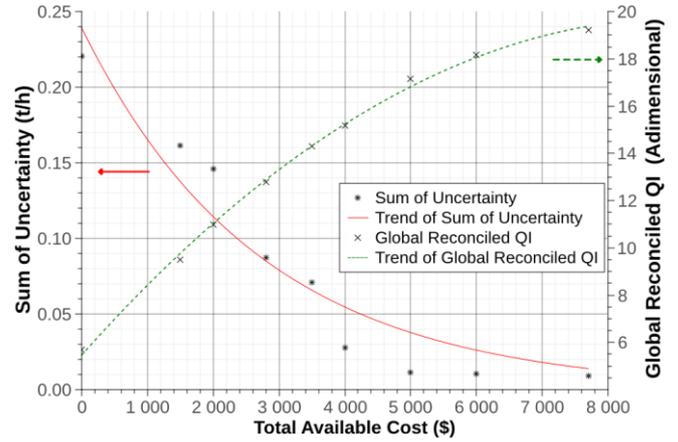


Fig. 3. Results of Example for Global QI MINLP Model

5. NOTATION

- a_i : sensibility factor of variable i (dimensionless);
- c_i : cost of a sensor i (US\$);
- c_T : total investment available for the sensor network (US\$);
- $f(z)$: mass or energy balances in plant;
- \mathbf{I} : the unit matrix;
- \mathbf{J} : the incidence matrix;
- M_w : list of variables on measurement system;
- M_p : list of parameters to control on measurement system;
- N_i : system node i ;
- q_i : binary variable which represents an installed sensor;
- QI_i : Quality of Information on the stream i (dimensionless);
- QI_G : Global Quality of Information of measurement system (dimensionless);
- S_i : water stream i ($t h^{-1}$);
- \mathbf{S} : matrix of variance in measure system ($t h^{-1}$)²;

\hat{S} : matrix of variance in reconciled measure system ($t h^{-1}$)²;

S_{QI} : matrix of Quality of Information in measured system (dimensionless);

\hat{S}_{ii} : the diagonal element i of \hat{S} matrix ($t h^{-1}$)²;

z_i : measured flow i ($t h^{-1}$);

Greek Letters

κ_i : proportionally constant in Quality of Information of variable i ;

σ_i : variance of variable i ($t h^{-1}$)²;

σ_i^* : threshold value for variance of variable i ($t h^{-1}$)²;

$\hat{\sigma}_i$: estimated variance of variable in measurement system ($t h^{-1}$)²;

ACKNOWLEDGMENTS

Authors gratefully acknowledge the financial support given by the Capes (Brazilian State Agency for Innovation), the ANP (Brazilian State Agency for Oil Regulation), the Fafen-BA (Operating Unit Nitrogen Fertilizer Plant in Bahia – Brazil) and the TECLIM-UFBA (Research Group at Clean Technologies) for their technical and financial support.

REFERENCES

- [1] S. Narasimhan, C. Jordache. *Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data*. 1st ed., p.350. Houston, TX: Gulf Professional Publishing. Retrieved from <http://www.sciencedirect.com/science/book/9780884152552>, 2000.
- [2] D. Fontana, R. Kalid, A. Kiperstok, M. Silva, E. Sales, J. Pacheco Filho, S. Oliveira, C. Perazzo. *Balanço Hídrico – uma nova sistemática*. XV Congresso Brasileiro de Engenharia Química. Anais... p.1-7. Retrieved from http://www.teclim.ufba.br/site/material_online/publicacoes/pub_art45.pdf, 2004.
- [3] M. A. F. Martins, C. A. Amaro, L. S. Souza, R. Kalid, A. Kiperstok. *New objective function for data reconciliation in water balance from industrial processes*. Journal of Cleaner Production, v. 18, n. 12, p. 1184-1189. doi: 10.1016/j.jclepro.2010.03.014, 2010.
- [4] M. Bagajewicz, M. Sánchez, *Sensor network design and upgrade for plant parameter estimation*. Computers & Chemical Engineering, v. 23, p. S593-S596. doi: 10.1016/S0098-1354(99)80146-X, 1999.
- [5] H. Tuy. *Convex programs with an additional reverse convex constraint*. Journal of Optimization Theory and Applications, v. 52, n. 3, p. 463-486. doi: 10.1007/BF00938217, 1987.
- [6] B. Weidema, M. Wesnas. *Data quality management for life cycle inventories—an example of using data quality indicators*. Journal of Cleaner Production, v. 4, n. 3-4, p. 167-174. doi: 10.1016/S0959-6526(96)00043-1, 1996.
- [7] R. P. Soares, A. R. SECCHI. *EMSO: A new environment for modelling, simulation and optimisation*. Computer Aided Chemical Engineering, v. 14, p. 947-952. Elsevier. doi: 10.1016/S1570-7946(03)80239-0, 2003.
- [8] P. Bonami, (Compiler), J. Lee, 2007. *Bonmin Users' Manual*. Available at: projects.coin-or.org/Bonmin/wiki/ (accessed 13.03.11.).