D.Brown[1], F.Maréchal[2], G.Heyen[3], and J.Paris[1]

# Data reconciliation and sampling protocol design, case of a paper deinking process

**Keywords**
Sensitivity matrix, sensor system, reconciled precision, key performance indicators, population based evolutionary algorithm, objective function, waste paper deinking mill.

## Abstract

An equation solver data reconciliation software has been used to build a validated model of a waste paper deinking mill, by combining control room measurements and process design specifications. An optimal sampling protocol to validate the model by using only control room measurements has been determined by identifying, with genetic algorithm programming, the additional sampling points and corresponding sensors required to compensate for the lack of redundant measurements.

## Introduction

Computer aided process simulation is an efficient design tool, which can help pulp and paper companies, pressured by global competition, and urged to comply with environmental regulations, to upgrade their facilities rapidly, and at low engineering and operating cost. Specifically, in the context of system closure, accurate models are needed to predict the impact of retrofit modifications on a given process. Data reconciliation is essential for process performance follow-up and simulation model calibration. Based on measurement redundancy, it is recommended as a preliminary step to process simulation. Data reconciliation has been extensively used in the petro-chemical and chemical industries, however only recently has it been applied to pulp and paper processes (Jacob and Paris, 2003). Due to the number of pieces of equipment and process streams involved, and in spite of the abundance of information that can be acquired using process sensors, additional measurements are often required to reach satisfactory levels of redundancy. This can be achieved at an acceptable cost by combining data reconciliation with equipment design specifications and process diagrams. The problem with this approach is the validity of these additional specifications, and their impact on the measurement corrections and precision. However, conducting systematic online process calibration by using data from a digital control system would ensure measurement precision and coherence. The benefits of real time mill data reconciliation have been discussed in detail by Heyen (2000). Jacob and Paris, (2003) have discussed the principles of reconciliation and the concept of measurement redundancy. They applied a heuristic approach based on local redundancy analysis of each process unit to plan sampling campaigns in an integrated TMP mill, to show how data reconciliation can improve reliability of simulations and help detect suspicious data. However even though computer aided, this flowsheet examination of a large scale process is cumbersome. To overcome this drawback, a method based on the sensitivity matrix analysis and the use of a genetic algorithm has been proposed and applied to an ammonia production plant by Heyen *et al* (2002) to automatically optimize the identification of appropriate additional sampling points. The first objective of this study was to calibrate a model of an old newspaper and magazine deinking mill by applying data reconciliation. This model was built by using various sources of information which trace back to different periods in time. The second objective was to design a sampling protocol to determine the minimum additional measurements required to enable process calibration by utilizing only control room measurements.

## Methodology

### Data reconciliation

In order to perform data reconciliation there must be an excess of information, beyond what is strictly needed to solve the system of equations that are used to model a process: this is redundancy. By subtracting the number of unmeasured variables from the number of model equations a global redundancy number can be determined:

$$R_G = p - m \tag{1}$$

Thus, it is possible to know if the problem can still be solved ($R_G \geq 0$) or if additional measurements are necessary ($R_G < 0$). From a mathematical standpoint data reconciliation is an optimization, carried out by means of an objective function that corresponds to the minimization of the weighted errors on measurements, under constraints representing the behavior of the process operations (mass and energy balances, separation rules, and thermodynamic behaviors). The data reconciliation problem can be expressed as:

$$\underset{X,Y}{Min}\, H = (Y - y)^T P (Y - y), \tag{2a}$$

under the constraint

$$F(X,Y) = 0, \tag{2b}$$

with,

$$F = AY + BX + C, \tag{2c}$$

and,

$$A_{ij} = \frac{\delta F_i}{\delta Y_j},\ B_{ij} = \frac{\delta F_i}{\delta X_j} \tag{2d}$$

If the assumption of statistically independent and normally distributed random errors on measurements (Madron, 1985) is made, P is the diagonal matrix of the inverses of the variances of the measured values. Therefore the objective function can also be expressed as:

$$\sum_{i=1}^{n} \frac{(Y_i - y_i)^2}{\sigma_i^2} \tag{3}$$

The constrained data reconciliation problem can be transformed into an unconstrained one by using the Lagrange formulation:

1. *D. Brown and J. Paris*, Département de génie chimique, École Polytechnique, Montréal, Canada. Correspondence should be addressed to: Jean Paris, Département de génie chimique, École Polytechnique, C.P. 6079, succ. Centre-ville, Montréal, (Québec) Canada H3C 3A7. Email: jean.paris@polymtl.ca
2. *F. Maréchal*, LENI, École Polytechnique Fédérale de Lausanne, Switzerland
3. *G. Heyen*, LASSC, Université de Liège, Belgium

$$\underset{X,Y,\lambda}{Min}\, L = (Y - y)^T P(Y - y) + 2\lambda F \quad (4)$$

Then, the following Euler equations are solved.

$$\frac{\delta L}{\delta Y_i} = 0 \quad i = 1, \ldots n \quad (5a)$$

$$\frac{\delta L}{\delta X_i} = 0 \quad i = 1, \ldots m \quad (5b)$$

$$\frac{\delta L}{\delta \lambda_i} = 0 \quad i = 1, \ldots p \quad (5c)$$

to yield the following system of equations:

$$PY + A^T \lambda = Py \quad (6a)$$

$$B^T \lambda = 0 \quad (6b)$$

$$AY + BX = -C \quad (6c)$$

The square matrix M and vectors V and D can then be defined such that:

$$M = \begin{bmatrix} P & 0 & A^T \\ 0 & 0 & B^T \\ A & B & 0 \end{bmatrix} \quad V = \begin{bmatrix} Y \\ X \\ \lambda \end{bmatrix} \quad D = \begin{bmatrix} Py \\ 0 \\ -C \end{bmatrix} \quad (7)$$

In this way, the solution of the validation problem can simply be expressed as:

$$V = M^{-1} D \quad (8)$$

Vectors X and Y are linear combinations of measured values y. The matrix $M^{-1}$ can be used to perform a sensitivity analysis i.e to evaluate to what extent the validated value of a variable is affected by other measured variables and their standard deviations (Heyen *et al*, 1996):

$$Y_i = \sum_{j=1}^{m+n+p} (M^{-1})_{ij} D_j \quad (9a)$$

$$Y_i = \sum_{j=1}^{n} (M^{-1})_{ij} P_{jj} y_j - \sum_{k=1}^{p} (M^{-1})_{i\,n+m+k} C_k \quad (9b)$$

The variance of a linear combination $Q_e$ of several variables $Z_e$ is calculated in the following way:

$$Q_e = \sum_{e=1}^{q} a_e Z_{ej} \quad (10)$$

$$\mathrm{var}(Q_e) = \sum_{e=1}^{q} a_e^2 \,\mathrm{var}(Z_e) \quad (11)$$

Thus the reconciled standard deviation of the reconciled measured variable $Y_i$ is given by:

$$\mathrm{var}(Y_i) = \sum_{j=1}^{n} \frac{(M_{i,j}^{-1})^2}{\sigma_j^2} \quad \forall i = 1, n \quad (12)$$

In practice, data reconciliation of a large scale process is an iterative procedure because a preliminary observation of available process data generally indicates that the problem of local redundancy is insufficient. Local redundancy i.e., the redundancy of the subsystems of streams in the vicinity of a process unit has been discussed by Jacob *et al* (2003). The redundancy of the system can also be determined by the form of the incidence matrices A and B defined in equation 2d.

**Analysis of redundancy**

The data reconciliation problem has an infinite number of solutions if the number of unmeasured variables is superior to the number of constraint equations. In this particular case, $R_G$ is negative and matrix B is rectangular and horizontal, and the sensitivity matrix is singular. If the number of unmeasured variables is just equal to the number of constraint equations, $R_G$ is equal to zero and B is a square matrix, and the solution of the problem can be obtained by considering all measured variables as constants; however, there can be no validation of process measurements. This is typical in process simulation. If the number of unmeasured variables is inferior to the number of constraint equations, $R_G$ is positive and B is rectangular and vertical, and measurements can be reconciled. For a solution to exist, matrix B cannot be horizontal. Furthermore, the sensitivity matrix cannot be singular. This happens for instance when the constraint equations are linearly dependant, when the measurements are not well distributed and certain parts of the process remain undetermined (i.e., when local redundancy is insufficient) or, when the constant variables are not well chosen and create an over specified problem. Structural analysis of the global incidence matrix [A B] (the matrix of measured and non-measured variables) can be applied to identify missing measurements, and measurements that cannot be validated, as well as the ones that can be corrected with redundant measurements (Kalitventzeff and Joris, 1987). It consists in outlining by line column permutation, two sub-matrices within both the measured and the unmeasured variables incidence matrices:

- For matrix A: a horizontal matrix of measurements that cannot be validated, and a vertical matrix of variables that can be recalculated and corrected by using the values of other measurements.
- For matrix B: a horizontal matrix of incalculable variables, and a vertical matrix of variables that can be calculated or validated.

If a set of incalculable variables exists, additional measurements are required. The question to be resolved then is where to locate the most favorable sampling points.

**Sampling protocol design**

The identification of additional sampling points can be done by using a method involving genetic algorithm programming (Heyen *et al,* 2002), that selects optimal sampling points for missing measurements. In this phase, a simplified formulation of the sensitivity matrix has been defined: the actual measurements have their own accuracy, while the unmeasured variables are considered as measured variables with a very high standard deviation ($10^{15}$):

$$\begin{bmatrix} Y \\ \lambda \end{bmatrix} = \begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix}^{-1} \begin{bmatrix} Py \\ -C \end{bmatrix} = M^{-1} \begin{bmatrix} Py \\ -C \end{bmatrix} \quad (13)$$

The different types of feasible measurements are defined in a database (e.g. flow rate, consistency, pressure, temperature, etc...). The measurement precision has to be specified, (fixed and/or variable errors). Other pertinent (but optional) information can be the cost and range of measuring devices. Heyen *et al* (2002) have emphasized the incentive of designing an online data reconciliation sensor system, in which cost weighting factors are used to minimize the investment required for such a system. Nonetheless, the optimization's objective function may also simply consist in the number of sampling points for a measurement campaign. Prior to optimization, the program verifies that reconcilable configuration can be found, assuming that all possible measurements are done. This provides an upper limit for the objective function. The use of a new measurement at a potential sampling point is obtained by assigning a 0-1 value to an integer variable.

$$\varepsilon_i^2 = \left[ \left( \sum_{g=1}^{s_g} \frac{\gamma_{i,g}}{\varepsilon_{meas\_typeg}^2} \right) + \frac{1}{\sigma_i^2} \right]^{-1} \quad \forall i = 1, n \quad (14)$$

The reconciled standard deviations are recalculated from the modified sensitivity matrix redefined by the sets of measurements selected by the algorithm. Then the objective function is computed by summing up the number of measurements (or sensors costs) and a penalty function constituted by the sum of the projected standard deviation of key performance indicators (or by default any process variable) resulting from the analysis of the sensitivity matrix.

$$\min_{\gamma_{f,g}}\left(\sum_{f=1}^{s_f}\sum_{g=1}^{s_g} Cost_g \cdot \gamma_{f,g} + \sum_{h=1}^{nKPI} Pen_h\right) \tag{15a}$$

with

$$\begin{cases} Pen_i = \dfrac{\operatorname{var}(Y_i)}{\operatorname{var}(Y_i)_{desired}} \quad if \quad \operatorname{var}(Y_i) \le \operatorname{var}(Y_i)_{desired} \\ Pen_i = 0.01\min\left(10, \dfrac{\operatorname{var}(Y_i)}{\operatorname{var}(Y_i)_{desired}}\right)^2 \quad if \quad \operatorname{var}(Y_i) > \operatorname{var}(Y_i)_{desired} \end{cases} \tag{15b}$$

Following Heyen *et al* (2002), two domains have been considered for the penalty function to retain solutions violating the constraint on the variance and thus maintain a broad spectrum of solutions from which an optimum can be identified by genetic algorithm search.

Data reconciliation, measurement redundancy analysis and optimal sampling point location have been applied to an old newspaper and magazine deinking process. This will now be discussed.

## Case study

The methodology was applied to a deinking mill located in Quebec. It uses 80% old newsprint (ONP) and 20% old magazines (OM) furnish to produce deinked pulp. The mill is located next to an integrated thermo-mechanical pulp and newsprint mill to which part of the deinked pulp is sent to produce 30% recycled content paper. The recycling facility was built in the early nineties, and was subsequently modernized at the end of the decade in order to increase its production capacity. During the upgrade, several modifications were made to the pulp treatment sequences and the process water circulation layout. It is estimated that the fresh water intake has been reduced from approximately 21 to 15 tons per odt of pulp produced. Fig. 1 shows a simplified layout of the present day mill. A detailed description of the process in its configuration at the time of the study is given in Brown (2004).

### Modelling assumptions for data reconciliation

The model of the mill has been developed by using the equation solver data reconciliation software VALI III (Belsim s.a, 2001). This software can perform data reconciliation and plant simulation and automatically transfers data between the two operations. The calibrated model has been built by using data from previous studies of the original mill (Walosik, 1999; Bonhiver et al, 1998; Savu et al, 2004), updated with specifications from control diagrams, and laboratory test benchmarks of the present day mill. Furthermore, several assumptions were made to reduce the number of required specifications. The density of cellulose was added to the component data base of VALI III and other physical properties of cellulose ignored since temperature variations are small in a deinking process. It was also assumed that waste paper enters the pulper at 15 °C, that the temperature of fresh water and whitewater entering the system is constant at 50 °C, that there is no heat loss from the process piping and equipment, and that output streams from units with multiple outlets are at equal temperature. Pressure variations in the piping network were accounted for when they were indicated by measurements from controller room printouts. Pressure drops were otherwise neglected.

In a preliminary step to data reconciliation, the VALI III solver determines the degree of global and local redundancy of the system. Five different types of specifications related to mass balances were used to compensate for the missing control room measurements:

- Two types of measured variables: consistency and flow rate. Generally, consistency and pressure variations are monitored using smaller time constants in control loops. Hence, consistency was preferred over flow rate when specifying dilution stream requirements from design specifications, the exception being when units such as pressure screens must operate at steady discharge.
- Three types of constant ratio specifications defined by link equations for separation units and stream splitters:
  - rejects or accepts mass ratios (expressed as the ratio between inlet and outlet mass flows on an oven dried basis),

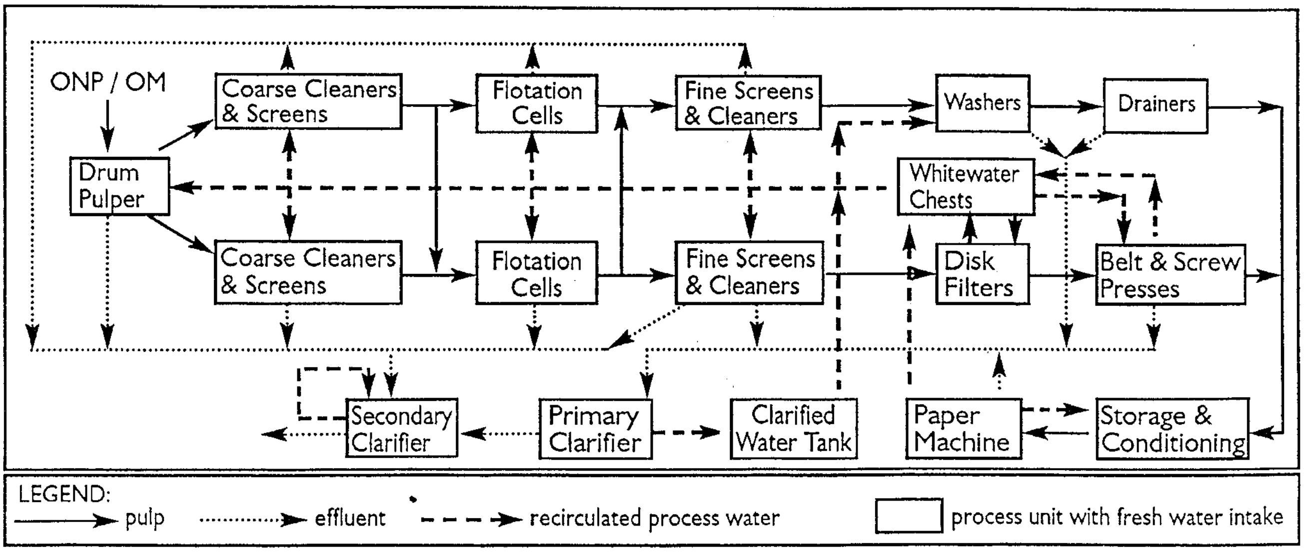


**Fig. 1. Process water flow diagram.**

- separation units' thickening ratios (expressed as the ratio of the outlet consistency to the inlet consistency),
- flow rate ratios for split streams.

In practice, only volumetric flow and consistency can be determined from the mill's control system, whereas specifications expressed as ratios are related to equipment performance parameters indicated on process diagrams, and cannot be directly measured during mill operation. These types of specifications are useful because they model process units, and thus confer the flexibility required to account for flow or consistency variations.

### Outline of sampling protocol design

After the initial data reconciliation step, the values of the precision of the measurements that were not obtained from the mill's control system were changed so as to redefine them as unmeasured variables. Process variables can be alternately defined by the VALI III software either as measured variables, unmeasured variables or constants simply by modifying the values of their precision.

The sampling protocol design was done for sections of the mill for which controller room printouts were available: the pulper and the second pulp treatment line. Cost weighting factors have been included to account for the possibility of designing an online sensor system; however this does not exclude using results for a measurement campaign. Annualized costs and sensors' precision are taken from the database of Gerkens (2002). These include pressure, temperature, flow, and consistency measuring devices. Two different cases have been considered:

- In the first case, ratio specifications are considered as constants; this implies that there is no variability in the separation ratios of the units or splitters of the process. The resulting measurement configuration would be capable of accounting for the overall process variability (e.g. the generally non-steady state behavior of the process due to sheet breaks, start-ups, slow downs, etc.).
- In the second case, ratio specifications are considered as variables, so that the performance of each individual separation unit may also be accounted for.

**Table 1. Process overall mass balances.**

| | Stream Description | Raw data Flow (kg/s) | Corrected by data reconciliation Flow (kg/s) |
|---|---|---|---|
| Inlets | Water | 77.20 | 78.14 |
| | Paper (ONP / OM) | 6.74 | 6.74 |
| | Whitewater | 63.45 | 63.57 |
| Total In | | 147.39 | 148.45 |
| Outlets | Sludge | 2.20 | 2.28 |
| | Other solid waste | 0.47 | 0.45 |
| | Effluent Water | 55.86 | 56.90 |
| | Pulp (10% Cs) | 18.88 | 18.42 |
| | Pulp (4.4% Cs) | 70.18 | 70.40 |
| Total Out | | 147.59 | 148.45 |

**Table 2. Measurements and reconciled measurements for selected streams.**

| Stream Description | Flow Rate (kg/s) | | | Consistency (%) | | |
|---|---|---|---|---|---|---|
| | Meas. | Rec. | D. (%) | Meas. | Rec. | D. (%) |
| Fresh water to whitewater tanks | 74.28 | 74.94 | 0.9 | 0 | 0 | 0 |
| Cloudy whitewater to pulper | 153.30 | 153.30 | 0.0 | 0.033 | 0.035 | 6.2 |
| **Upper Cleaning Line (L1)** | | | | | | |
| Cloudy whitewater to coarse cleaners | 50.16 | 53.92 | 7.5 | 0.033 | 0.035 | 6.2 |
| Whitewater to flotation cells | 83.97 | 78.25 | 6.8 | 0.095 | 0.096 | 1.7 |
| Clarified water to fine cleaners | 98.15 | 99.76 | 1.6 | 0.034 | 0.036 | 7.0 |
| Clarified water to washers | 73.86 | 73.94 | 0.1 | " | " | " |
| **Lower Cleaning Line (L2)** | | | | | | |
| Acid WW to primary coarse screen | 8.27 | 8.25 | 0.3 | 0.035 | 0.036 | 2.1 |
| Acid WW to secondary coarse screen | 2.35 | 0.22 | 90.6 | " | " | " |
| Acid WW to secondary fine screen | 53.35 | 53.65 | 0.6 | " | " | " |
| Acid WW to tertiary fine screen | 31.57 | 31.95 | 1.2 | " | " | " |
| Acid WW to secondary fine cleaners | 70.07 | 72.93 | 4.1 | " | " | " |
| Acid WW to tertiary fine cleaners | 26.37 | 27.29 | 3.5 | " | " | " |
| Acid WW to quaternary fine cleaners | 6.01 | 5.75 | 4.2 | " | " | " |
| Acid WW to reverse fine cleaners | 14.03 | 14.01 | 0.2 | " | " | " |
| # 1 disk filter cloudy water | 113.30 | 113.24 | 0.1 | 0.04 | 0.04 | 0.0 |
| # 2 disk filter cloudy water | 159.14 | 159.59 | 0.3 | 0.025 | 0.025 | 0.0 |
| # 1 disk filter clear water | 113.30 | 113.22 | 0.1 | 0.03 | 0.03 | 0.0 |
| # 2 disk filter clear water | 159.14 | 159.57 | 0.3 | 0.025 | 0.025 | 0.0 |
| Filtered whitewater to belt press | 9.93 | 9.93 | 0.0 | 0.005 | 0.005 | 0.0 |
| Filtered whitewater to disk filter # 1 | 39.71 | 39.71 | 0.0 | " | " | " |
| Filtered whitewater to disk filter # 2 | 38.06 | 38.02 | 0.1 | " | " | " |

Notes. Meas.: measured value; Rec.: reconciled value; WW: whitewater; D.: difference between reconciled and measured value

**Table 3. Summary of the analysis of the incidence matrix.**

| Total number of... | Constant ratio specifications | Variable ratio specifications |
|---|---|---|
| Equations | 405 | 405 |
| Measured variables | 43 | 43 |
| Unmeasured variables | 448 | 486 |
| Constants | 130 | 92 |
| Equations with no influence on the validation problem | 40 | 20 |
| Measurements that cannot be validated | 21 | 21 |
| Unmeasured variables that cannot be validated nor calculated | 408 | 466 |
| Additional measurements required (minimum) | 45 | 90 |
| Global redundancy number ($R_G$ = p-m) | -43 | -81 |

## Results and discussion

### Data reconciliation

Table 1 compares global mass balances for the entire process computed with raw data, to the same balances obtained by data reconciliation. Differences can be noted for the unmeasured flow variables (only the ONP and OM feed was specified as a measurement). It was deemed preferable to consider the process diagram specifications as indicative values, because when compared to design specifications, control room printouts indicated significant variability. Also, production reports from the mill show that the total pulp output may vary considerably from day to day. In this respect, a real time simulation model would have to be flexible enough to account for process operation variability. Table 2 shows selected measured and reconciled values of flow rates and consistency for some major process water streams. Some variables are significantly corrected

Table 4. Summary of the sampling protocol optimization (with weighting factors for a sensor system design).

| Maximum number of possible measurements | Upper limit of objective function (CAN $) |
|---|---|
| 385 | 354,482 |

| Best solution after 24 020 evaluations | Constant RS[a] | Variable RS[a] |
|---|---|---|
| Additional sensors required | 47 | 97 |
| Objective function (CAN $) | 35,511 | 84,340 |

| Measurement ()[b] | Weighting[c] (CAN $) | Precision[d] | Number of Measurements Present | Constant RS[a] | Variable RS[a] |
|---|---|---|---|---|---|
| Temperature (1) | 824 | 1.5° C | 0 | 11 | 19 |
| Pressure (1) | 808 | 1.5% | 30 | 50 | 56 |
| Volumetric flow (2) | 904 & 1, 585 | 1.5% | 8 | 16 | 30 |
| Consistency (6) | 793 to 1, 585 | 1% | 5 | 13 | 36 |
| **Total** | - | - | **43** | **90** | **141** |

Notes. a: RS = Ratio specifications; b: number of sensor types per type of measurement; c: weighting factors: annualized costs; d: the error is absolute for temperature measurements, and relative for other measurements

while in other cases there is no correction at all. A gross error (90% difference) has been detected for the flow of whitewater to the secondary coarse cleaners of the lower pulp treatment line; the reconciled data is otherwise within close range of the measured values (the second largest difference is 7.5%). It should be stressed though, that data is not necessarily accurate because it is not corrected. Lack of local redundancy (Jacob and Paris, 2003), causes certain measurements to be corrected only by their own value, which means that they are not reconciled. The currently calibrated model was built with data from various sources and periods in time, but for the oldest parts of the mill, the operating conditions are the same as before. It can be argued that the reconciled model is still a valid representation of the process. However, the design of a new sampling protocol was of interest, to validate it with additional actual data.

### Sampling protocol design

Table 3 presents an overview of the analysis of the incidence matrix of the system, after measurements not provided by the control system have been removed from the reconciled model, and for both modeling assumptions previously stated in the outline of the sampling protocol design. The summary of the sampling protocol design is presented in Table 4. In the actual configuration, 43 control room measurements are available, but from Table 3 it has been shown that they do not suffice to allow for the recalculation and validation of all other process variables. The results show that:

- the constant ratio specification configuration would require doubling the number of measurements, which is close to four times less than the maximum number.
- the variable ratio specification configuration would require tripling the number of measurements, which is close to three times less than the maximum number.

Although it is never certain that global optimality is reached when genetic algorithm programming is used and although the total number of required measurements might still have been reduced by weighting all measurements equally, the results obtained would nevertheless allow effective data reconciliation. The interest of the software is in automated measurement campaign planning, reducing the amount of unnecessary measurements, and avoiding using fastidious heuristic methods.

The added incentive of the design of a retrofitted sensor system would be to provide a monitoring tool to contribute to rapid detection of system defaults (decalibrated sensors, process unit breakdowns or leaks), hence to minimising losses and obtaining actualized mill balances for accounting and process follow-up (Heyen, 2000). This is the only way to ensure that the results of a simulation model rigorously correspond to the output of a process in operation.

## Conclusions

Data reconciliation has been applied to build a validated model of a recycled paper deinking plant. Controller measurements were combined with equipment and process design specifications, to compensate for missing data. Five different types of specifications related to mass balances were added to the model: consistency, flow rate, and three types of ratio specifications for rejects or accepts rates, thickening ratio and stream splitters. Despite the fact that the inputs to the process may vary greatly, the model is still generally consistent because ratio design specifications allow accounting for measurement variability.

The key to effective reconciliation is measurement redundancy. If there is not sufficient information, data cannot be reconciled. In the currently calibrated model, a satisfactory level of redundancy has been attained with the use of data from various sources. In the perspective of eventually updating the data from the simulation, measurement specifications from sources other than the control system were eliminated and a method based on genetic algorithm programming was applied to optimize a sampling protocol. This optimization program could also be used to design a sensor system in order to conduct online data reconciliation by using control room measurements only. Depending on whether or not the variability of separation units and stream splitters is taken into account, the sampling campaign would require doubling or tripling the number of measurements. The computer generated solution has the advantage of avoiding fastidious planning by local redundancy observation.

## Notation

### Roman characters

| | |
|---|---|
| A: | Jacobian matrix (incidence matrix) for measured variables (dimension n, p) |
| B: | Jacobian matrix (incidence matrix) for unmeasured variables (dimension m, p) |
| C: | matrix of constants |
| $cost_g$: | cost of measurement of type g |
| F(X, Y): | vector of process equations |
| H: | objective function in Lagragian formulation |
| n: | number of measured variables |
| nKPI: | number of key performance indicators |
| m: | number of unmeasured variables |
| L: | objective function with integrated constraints |
| M: | Jacobian matrix of the system |
| p: | number of equations |
| P: | inverse variance-covariance matrix of measured variables (dimension n, n) |
| $Pen_h$ : | penalty function of KPI h |
| q: | number of elements in linear combination Q |
| $Q_z$: | linear combination of variables Z |
| $R_G$: | global redundancy number |
| $s_f$: | number of possible additional measurements |
| $s_g$: | number of measurement types |

var ($Y_i$): reconciled variance of reconciled measured variable i
X: vector of unmeasured variables
y: vector of measured variables
Y: vector of reconciled measured variables

**Greek characters**

$\gamma_{f,g}$: binary integer variable $\in\{0,1\}$
$\varepsilon_i$: reconciled standard deviation of reconciled measured variable i
$\varepsilon_{meas\_type\ g}$: precision of sampling measurement type g
$\lambda$: Lagrange multipliers
$\sigma_i$: standard deviation of measured variable i

**Indices**

e: linear combination example
f: possible additional measurements
g: measurement types
h: key performance indicators.
i: measured values
j: unmeasured values
k: constants

*Exponents*

T: transposed matrices


## Acknowledgements

The authors would like to thank MRN Québec and the EU-Canada Cooperation Agreement on Higher Education and Training, Academic Mobility Program, from HRD Canada that have contributed to funding this project, as well as the mill's staff for providing time and information towards the realization of the study. This work was done on the behalf of the Chair in Process Integration in the Pulp and Paper Industry at École Polytechnique. The support of the chair is gratefully acknowledged.


## References


Belsim, *VALI 3 User Guide*, Belsim s.a, St-Georges-sur-Meuse, Belgium, 1, 8 (2001).

*Bonhivers, J.-C., S. Belon-Gagnon, J. Paris* : Simulation dynamique de l'atelier de désencrage de l'usine de Kruger à Bromptonville, Québec. Manual de reference, École Polytechnique, Montréal, Canada (1998).

*Brown, D.* : Analyse de la fermeture de circuits d'eau d'un procédé de désencrage de vieux papiers par la programmation non linéaire mixte en variables entières. Mémoire de M.Sc.A., École Polytechnique, Montréal (2004).

*Gerkens, C.* : Conception rationnelle de systèmes de mesure dans les procédés chimiques. Rapport de projet, Université de Liège, Liège, Belgium (2002).

*Heyen G., E. Maréchal, B. Kalitventzeff*: Sensitivity Calculations and Variance Analysis in Plant Measurement Reconciliation. Computers and Chemical Engineering, 20S, 539–44 (1996).

*Heyen, G.*: Application of Data Reconciliation to Process Monitoring. Symposium ESCAPE 2000, Carthagena de Indias, Colombia (2000).

*Heyen, G., M.N. Dumont, B. Kalitventzeff*: Computer-Aided Design of Redundant Sensor Networks. ESCAPE 12, The Hague, Netherlands, J. Grievink and J. van Schyndel, Elsevier, 691–5 (2002).

*Jacob, J., J. Paris*: Data Sampling and Reconciliation, Application to Pulp and Paper Mills. Appita Journal, Part I: Methodology and Implementation, 56 (1), 25–9&52, (2003), Part II: Case Studies, 56 (2), 116–21 (2003).

*Kalitventzeff, B., P. Jori.*: Process Measurements Analysis and Validation. Proceed. CEF87, The Use of Computers in Chemical Engineering, Taormina, Italy, 41–46 (1987).

*Madron, F.*: A New Approach to the Identification of Gross Errors in Chemical Engineering Measurements. Chem. Eng. Sci., 40 (10), 1855–1860, 1985.

*Savu, E., S. Sarailh, F. Maréchal, J. Paris* : Fermeture des circuits dans un procédé de désencrage. Analyse et simulation P&P Can., 105 (12), 28–33 (2004).

*Walosik, S.* : Gestion de l'eau dans le procédé de désencrage : problématique et étude de cas. Mémoire de M. Sc. A., École Polytechnique, Montréal, Canada (1999).



Received December 1, 2003.