# STRATEGY FOR IMPROVING DATA QUALITY FOR A KRAFT PULP MILL RECAUSTICIZING PLANT

Taiwen Jiang (1), Bingzhen Chen (2), Khalid Jasim (3) & Paul Stuart (1)*

(1) École Polytechnique, Montréal, Québec H3C 3A7 (CANADA)
(2) Tsinghua University, Beijing 100084 (CHINA)
(3) Domtar Papers Limited, Cornwall, Ontario K6H 5S3 (CANADA)

*Abstract*

Steady state simulation of a kraft pulp mill recausticizing plant was completed for the purposes of investigating process changes for an area upgrade. Due to errors in the measured data as well as assumptions made due to a lack of necessary information, the initially completed simulation was not the best representation of the process at steady state. A comprehensive approach has been used to improve and validate the simulation. Firstly, a multi-scale data processing technique based on wavelet transform was used to eliminate errors caused by temporal events, including random noise and abnormalities in the raw measurements. Steady state was then detected using the wavelet-based method. Data reconciliation was used to reduce systematic errors and to calculate some of the necessary variables that were not measured. Combined with process knowledge, the simulation was validated using the reconciled data.

*Keywords*

Recausticizing plant, data processing and reconciliation, steady state detection, Wavelet Transform.

**Introduction**

An upgrade of the recausticizing plant of a pulp mill was proposed with the goals of process stabilization and general improvement of process operations. In this context, a validated steady state simulation was required for providing reliable recommendations, given limited data availability and quality.

To calibrate and validate the simulation, real-time data were collected from the plant data management system and other sources. However, the measured process data inherently contained various errors, including random errors caused by random or temporal events and systematic errors (gross errors) caused by instrument malfunction or process leaks. Besides true measurement errors, departure from steady-state operations also causes apparent violation of conservation laws. Furthermore, some necessary variables are not measured due to sampling restrictions, measuring technique, or instrument failure.

A comprehensive strategy for improving the data quality is proposed integrating data process and reconciliation techniques (Figure 1): (1) Data processing reduces random errors and extracts the process trends. (2) Steady state is detected and the processed data are consequently selected from the identified steady-state durations. (3) Reconciliation techniques are then used to eliminate the measurement conflicts and to improve the consistency and completeness.

* To whom all correspondence should be addressed: Paul.Stuart @polymtl.ca

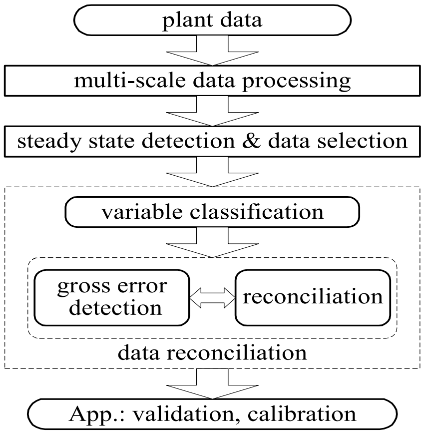


*Figure 1. Strategy for improving data quality.*

**Multi-scale Data Processing Using Wavelet Transform**

A wavelet-based multiscale data processing approach was proposed by Jiang *et al.* (2000) to eliminate temporary errors, including random noise and abnormalities, and to extract process trends. The process measurements are hierarchically cut into different frequency components at different scales by multi-scale Wavelet Transform. At each scale, the outline of the process is represented by the so-called smoothed signal, while the variation is represented by the detailed signal.

Measurement noise has frequencies distinctly higher than those of the process variation, hence is cut into the fine scales and subsequently reduced by thresholding the wavelet modulus.

An abnormality is typically referred to as a spike in measurements with a supernormal peak value and a short duration and can be identified by a coupling of wavelet modulus maxima with opposite sign and large amplitudes.

**Steady State Detection Using Wavelet Transform**

An effective method using wavelet transform technique for steady state detection was presented by Jiang (2000).

The first-order Wavelet Transform $Wf(t)$ of a signal $f(t)$ is proportional to its first derivative (Mallat and Zhong, 1992). Hence the wavelet modulus $Wf$(t) measures the variation of the signal. The signal is said to be under steady status when the modulus is equal or close to 0, as duration $(0, t_1)$ in Figure 2. Meanwhile inflexion points, peak-value or valley-value points of the signal, as duration $[t_2, t_3]$, which also present as zero points of $Wf$(t), are distinguished from the true steady points according to the non-zero modulus of the second-order Wavelet Transform $WWf(t)$, which is accomplished by performing the Wavelet Transform on $Wf(t)$.

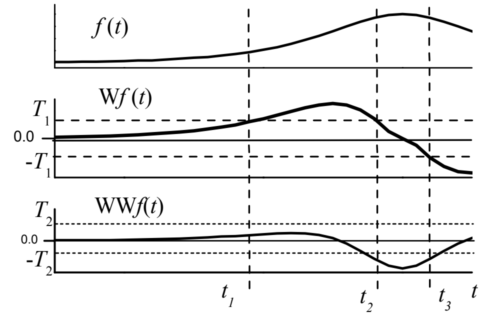


*Figure 2. Wavelet based steady state detection.*

A *steady state index*, $0 \leq \beta \leq 1.0$, calculated according to the wavelet modulus, is used to measure the degree of steady state: $\beta = 0$ indicates the unstable status and $\beta = 1.0$ indicates that the steady state is achieved.

**Data Reconciliation Using Sigmafine**

Data reconciliation is the procedure of optimally adjusting measured data so that the adjusted values obey the conservation laws and other process constraints, from which unmeasured variables can also be calculated. Refer to Crowe (1996) and Romagnoli (2000) for additional background.

*Statement of steady state data reconciliation*

Data reconciliation can be mathematically defined as a problem of constrained least-square errors between measured and adjusted values:

$$\begin{aligned} &\text{Min} \quad (\mathbf{x}_1 - \mathbf{x}_0)^{\mathrm{T}} \mathbf{W} (\mathbf{x}_1 - \mathbf{x}_0) \\ &\text{s.t.} \qquad f(\mathbf{x}_1, \mathbf{u}) = \mathbf{0} \end{aligned} \tag{1}$$

where $\mathbf{x}_0$, $\mathbf{x}_1$ are respectively the vectors of measured values and adjusted (reconciled) values, $\mathbf{u}$ is the vector of unmeasured variables, and $\mathbf{W}$ is a weighting matrix. The function $f$ represents the constraints, can be a linear, bilinear or nonlinear set of formulations.

*Data reconciliation using Sigmafine*

*Simafine* from OSI is a commercial software for steady state data reconciliation, which is well integrated with OSI PI. Both mass balances and component balances are included as constraints in our work, making it a bilinear problem. The constraints at the unit $k$ are written as

$$\begin{aligned} &\sum_j A_{j,k} F_j = 0 \\ &\sum_j A_{j,k} F_j C_{j,i} = 0 \end{aligned} \tag{2}$$

where $F_j$ is the total flow of stream $j$, $C_{j,i}$ is the mass fraction of component $i$ in stream $j$, and **A** is the incidence matrix representing the plant structure.

The technique employed by Sigmafine to solve the reconciliation problem is based on Kalman filtering and does not require matrix inversion. Sigmafine can handle large-scale system rapidly without convergence problem, which is particularly practical for industry application. Sigmafine allows the measurement test method and the constraint test method to detect gross errors in individual measurements and around units, respectively.

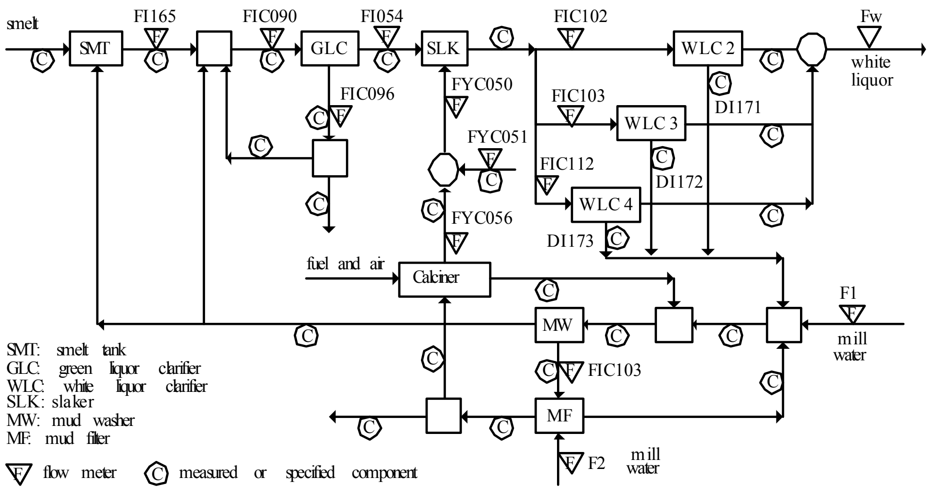


*Figure 3. Flow sheet of a kraft pulp mill recausticizing plant.*

**Application to Recausticizing Plant**

The recaustiucizing plant flow sheet is shown in Figure 4. In order to validate the simulation, real-time data were collected from the plant. The duration of the investigated data set was 6 days in March 2001 with sampling time interval of 5 minutes.

*Data processing*

Measurements of 25 key process variables were processed using the method described above. Figures 4 (a) and (b) show the treatment of two variables. It can be seen that both random noise and abnormalities in the measurements are eliminated effectively, and the process trends extracted without excessive distortion.

*Steady state detection and data set selection*

Five critical variables, covering flow rate, temperature and composition, were used to measure the process status. After the process steady state index is computed and rounded, eight different steady-state durations were identified, each of which indicated one scenario of the operation. Finally eight sets of steady-state data were selected from the corresponding identified duration. Figure 4(c) shows the identified steady-state durations.

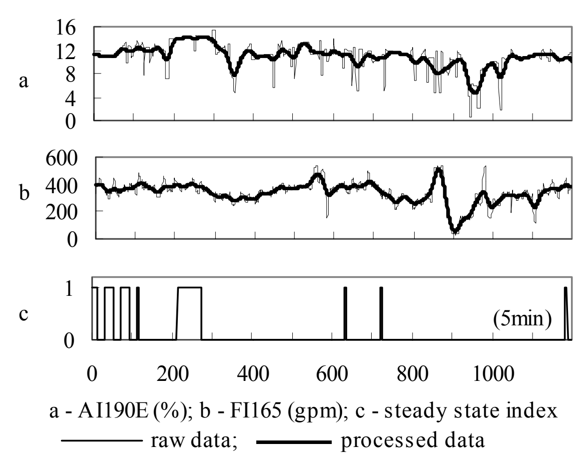


*Figure 4. Data processing and steady state detection using wavelet-based method.*

*Data reconciliation*

Sigmafine 3.0 was used to perform the data reconciliation. The following parameters were specified according to historical data or plant engineer experience: density and mass fraction of components in each stream and tolerance of each meter. For the purpose of practical reconciliation, a bi-component system was defined with the components: liquor and dregs. There were twelve flow rate and three component variables measured and two flow rate and 22 component variables specified. According to the solvability analysis, the variables around the calciner were neither reconcilable nor estimable, while the slaker was identified only mass balance applicable since a chemical reaction occurs inside. Each of the eight selected data sets were then reconciled. Column *A* and *B* of Table 1 respectively give the processed measurements and the reconciliation results of seven key variables in one set. Gross error was detected in the flow rate (FIC096) of underflow from the Green Liquor Clarifier, and was corrected. The total flow rate of Clarified White Liquor, *F*w, is estimated based on the solid concentrations in mud flows, which are measured by DI171, DI172 and DI173. Operating problems of the White Liquor Clarifiers in the plant caused significant departure of the solid concentration measurements from those under normal operating conditions. As a result, the estimation of *F*w (marked with superscript $\alpha$) according to the measurements is significantly different from normal conditions (marked with superscript $\beta$).

*Table 1. Reconciliation and validation re flow rate variables*

| Variables | A | B | C |
|---|---|---|---|
| FI165 (gpm) | 347.5 | 316.4 | 318.8 |
| FIC090 (gpm) | 318.3 | 321.6 | 325.1 |
| FIC054 (gpm) | 317.2 | 314.8 | 318.8 |
| FIC096 (gpm) | 14.9 | 6.8 | 6.2 |
| FYC050 ($ft^3$ /min) | 2.31 | 2.34 | 2.47 |
| FIC102+FIC103+FIC112 | 278.0 | 281.1 | 318.4 |
| *F*w (gpm) | | $81.4^{\alpha}$<br>$194.4^{\beta}$ | $218.6^{\beta}$ |

$\alpha$: abnormal operating conditions
DI171 = 16.4%; DI172 = 17.5%; DI173 = 16.6%
$\beta$: normal operating conditions
DI171, DI172, DI173 = 35%

*Validation of the steady state simulation*

The steady state simulation of the recausticizing plant was initially completed using WinGEMS from Pacific Simulation. The flow rate variables of simulation (Column *C* in Table 1) were validated using the reconciled data (Column *B* in Table 1). It should be noted that *F*w, the flow rate of Clarified White Liquor, was validated using the estimated value according to normal operating conditions instead of the true measurements since significant departure from normal operation exists. For additional validation, the chemical composition variables were compared with the processed measured data (no reconciliation available) and the data from a mass balance spreadsheet done by an engineer on-site (Column *C*, *B* and *A* in Table 2 respectively). It can be seen that the simulation is reasonable and acceptable for subsequent analysis.

*Table 2. Validation of chemical composition*

| Variables | A | B | C |
|---|---|---|---|
| GL NaOH (g/l $Na_2O$) | 19.6 | 30.6 | 17.0 |
| GL TTA (g/l $Na_2O$) | 125.4 | 120.6 | 121.1 |
| GL Sulphidity (%) | 10.9 | 11.0 | 13.8 |
| WL NaOH (g/l $Na_2O$) | 84.7 | 95.7 | 92.1 |
| WL TTA (g/l $Na_2O$) | 115.1 | 130.1 | 129.6 |
| WL Sulphidity (%) | 11.9 | 11.0 | 14.3 |

GL: green liquor; WL: white liquor

**Conclusions**

Real-time data from a kraft pulp mill recausticizing plant was treated through a comprehensive approach integrating data processing techniques and reconciliation techniques. The steady state simulation was then validated using the processed data and reconciled data. The validation results show that the simulation is reasonable for practical applications. The simulation was used to carry out a process sensitivity analysis, an investigation of pulp production increase, and to provide insight into capital effective options for recausticizing plant upgrade alternatives. Based on the analysis, several upgrade recommendations have been made.

In the future, the simulation could be linked to the PI data management system in order to stabilize the mill variability. Future tasks should focus on measuring the true variability of mill streams, and identifying methods to reduce this variability.

**References**

Crowe, C. M. (1996). Data reconciliation - progress and challenges. *J. Proc. Cont.*, **6**, 89-98

Jiang, T. W. (2000). PhD thesis: Wavelet Transform and its applications to soft-sensor of oil product qualities. Beijing: Tsinghua University.

Jiang, T. W., B. Z. Chen and X. R. He (2000). Industrial application of Wavelet Transform to the on-line prediction of side draw qualities of crude unit. *Computers and chemical engineering*, **24**(2-7), 507-512.

Mallat, S. and S. Zhong (1992). Characterization of Signals from Multiscale Edges. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **14**(7), 710-732.

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