Statistics & Control, Inc.

White Paper

OptiRamp® Material Balance Reconciliation

A Statistical Approach to Data Reconciliation for Continuous Processes

Vadim Shapiro
Dmitriy Khots, Ph.D.

Table of Contents

Introduction

An integral part of any industrial process is the material and energy balance. The notion of material balance comes from the physical law of conservation of mass that can be simply summarized by "what comes in must come out." Similarly, energy balance is derived from the first law of thermodynamics, which states that energy cannot be created or destroyed. Establishing material and energy balances assures process managers that there are no costly leaks or equipment failures within the process. The *OptiRamp* Material Balance Reconciliation (MBR) Submodule provides real-time reconciliation of material and energy balances by identifying measurement errors that could occur due to random noise, equipment failure, or unexpected leaks.

The MBR uses sophisticated statistical and optimization techniques to minimize the discrepancies within the existing material and energy balance equations for both local streams and global plant production. The submodule's algorithms are dynamic in nature: material balance reconciliation is performed for both steady state and transitional state processes. For the latter, genetic algorithms (described in the *OptiRamp* Optimized Scheduling Submodule white paper) estimate the process transition times.

The *OptiRamp* MBR Submodule performs the following functions:

- Detects faulty measurement signals
- Adjusts material and energy balance
- Estimates measurements and calculation errors
- Identifies faulty material movements—the system alerts the user whenever streams or unexpected leaks attain predefined limits
- Balances the transition state stream
- Archives algorithm execution data used for monitoring or by other *OptiRamp* modules and submodules

The MBR Submodule is fully integrated with the *OptiRamp* system. The *OptiRamp* Real-Time Optimization Submodule provides the MBR with the full suite of optimization tools. This integration allows MBR to reconcile material and energy balances and to provide operator recommendations in real time, which is an invaluable resource for Statistics & Control, Inc., (S&C) customers.

Balance Model

Material and energy balances are derived from the laws of conservation. The idea is that mass or energy cannot be created or destroyed within the confines of an isolated physical system. Typically, the industry standard for an isolated physical system is either a production unit or an entire plant. The material balance of a system can be generally characterized by the relationship $Input = Output + Accumulation$.

For energy balance, output can include losses to the environment. Preparing material or energy balances is outside the scope of this paper, but the following are general best practice guidelines:

- Draft the overall material and energy balance (e.g., using Sankey diagram shown in Figure 1)
- Use simple segmentation of the overall system
- Minimize the number of input/output streams within the system
- Include recycle streams
- Standardize all measurement units
- Include start-up/maintenance/shutdown resource consumption
- Use accurate process signal measurements
- Account for multi-level material and energy balances from the overall plant to each production unit—balances from each unit should be in agreement with the overall balance

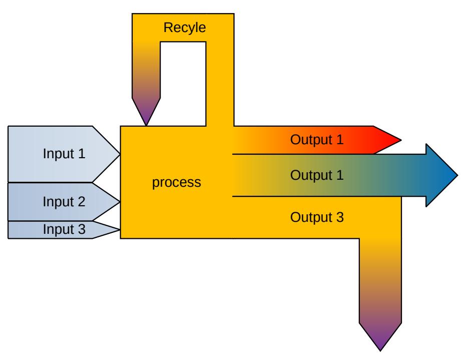

Figure 1. Sankey diagram of energy flows

The following discussion can be applied to either material or energy balances. For steady state processes with multiple flows, let $X_i(t)$, $i = 1,...,n$ denote the inflows; $Y_j(t)$, $j = 1,...,m$ denote the outflows; and $Z_p(t)$, $p = 1,...,k$ denote the accumulated quantities within each flow infrastructure. The material or energy balance is then described by equation (1):

$$\sum_{i=1}^{n} X_i(t) = \sum_{j=1}^{m} Y_j(t) + \sum_{p=1}^{k} Z_p(t) \qquad (1)$$

Transition state processes with multiple flows require as additional term: τ_j, $j = 1,\ldots,m$ represents the lag for each outflow. The material or energy balance is then described by equation (2):

$$\sum_{i=1}^{n} X_i(t) = \sum_{j=1}^{m} Y_j(t+\tau_j) + \sum_{p=1}^{k} Z_p(t), \quad (2)$$

where t ranges over a pre-defined timeframe $t \in [T_1, T_2]$. The *OptiRamp* MBR Submodule becomes an extremely powerful tool for business users whenever equations (1) or (2) are out of synch.

Balance Reconciliation

One of the main concerns of process operators occurs whenever there is a discrepancy within established material or energy balances. The primary causes for such discrepancies are signal noise, equipment failure, and unknown leaks. The MBR Submodule eliminates random measurement signal errors via data reconciliation techniques and identifies equipment failures/leaks (gross errors) by utilizing statistical hypothesis testing techniques and process models constructed by the *OptiRamp* Modeling Submodule.

Data Reconciliation

The *OptiRamp* MBR Submodule transforms material balance equations (1) and (2) into an optimization problem and minimizes the discrepancies by finding the optimal signal errors. *The key assumption for data reconciliation is that there are no gross errors in the system.* The next section discusses gross error detection.

Given a set of inflows $X_i(t)$, $i = 1,\ldots,n$, outflows $Y_j(t)$, $j = 1,\ldots,m$, and accumulations $Z_p(t)$, $p = 1,\ldots,k$, let $\delta_{X_i}, \delta_{Y_j}, \delta_{Z_p}$ denote the tolerances for inflows, outflows, and accumulations, respectively. The tolerances are individually chosen for each flow measurement based on measurement device characteristics and confidence levels dictated by the process. Tolerances can also be defined for certain calculated variables. For example, to calculate gas accumulation within a plant's fuel network, a series of assumptions are made with respect to the fuel network volume and fuel gas composition; thus, tolerances can be defined according to the calculation's confidence level.

Material and energy balance reconciliation occurs either globally across the entire plant or at each individual production unit. The measured and calculated variable errors are estimated within the acceptable tolerances. The statistical data reconciliation method is based on least squares optimization. Let Δ_{X_i}, Δ_{Y_j}, Δ_{Z_p} denote the target errors in $X_i(t)$, $Y_j(t)$, and $Z_p(t)$, respectively. The optimization problem for steady state processes is then set up as shown in equation (3)

$$\sum_{i,j,p}\left[\left(X_i(t)+\Delta_{X_i}\right)-\left(\left(Y_j(t)+\Delta_{Y_j}\right)+\left(Z_p(t)+\Delta_{Z_p}\right)\right)\right]^2, \quad (3)$$

minimizing the equation subject to constraints $\left|\Delta_{X_i}\right|\le\delta_{X_i}$, $\left|\Delta_{Y_j}\right|\le\delta_{Y_j}$, and $\left|\Delta_{Z_p}\right|\le\delta_{Z_p}$. For transition state processes, the optimization problem is given with similar constraints, but the objective function is given by equation (4).

$$\sum_{i,j,p}\left[\left(X_i(t)+\Delta_{X_i}\right)-\left(\left(Y_j(t+\tau_j)+\Delta_{Y_j}\right)+\left(Z_p(t)+\Delta_{Z_p}\right)\right)\right]^2 \quad (4)$$

Defining each outflow lag, τ_j, occurs during the process modeling stage described in the *OptiRamp* Modeling Submodule white paper. Its values are used in this optimization problem as known quantities. The *OptiRamp* Real-Time Optimization Submodule routines are used to find the optimal solution to the stated problem. The solution for the optimization problem is the set of estimated measurement/calculation errors $\bar{\Delta}_{X_i}$, $\bar{\Delta}_{Y_j}$, and $\bar{\Delta}_{Z_p}$ as well as the balanced measurement/calculated signals computed via $\bar{X}_i(t)=X_i(t)+\bar{\Delta}_{X_i}$, $\bar{Y}_j(t)=Y_j(t)+\bar{\Delta}_{Y_j}$, and $\bar{Z}_p(t)=Z_p(t)+\bar{\Delta}_{Z_p}$.

Gross Error Detection

The MBR Submodule detects four types of gross errors: bias, complete failure, drift, and precision degradation, as shown in Figure 2.

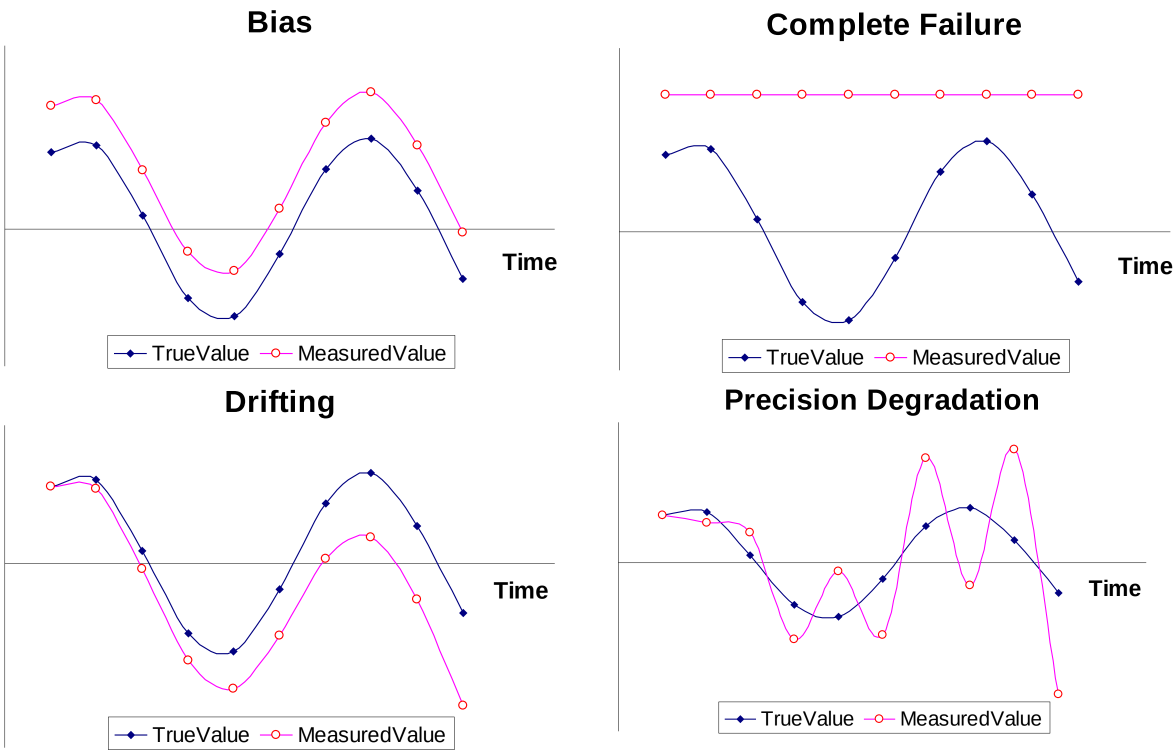

Figure 2. Gross error types

MBR uses three methods for gross error detection:

- Hypothesis testing using the Generalized Likelihood Ratio (GLR) test
- Tjao-Biegler test
- Neural network test

In the GLR test, the null hypothesis (H_0) that no gross errors exist (i.e., the expected value of the material or energy balance residual given by equation [5] is zero) is tested against the alternative hypothesis (H_1) that either a process leak or a measurement bias is present (i.e., the expected value of the material or energy balance residual is dependent on the unknown bias size across any of the measurement channels $X_i(t)$, $Y_j(t)$, or $Z_p(t)$, as shown in equation [6]).

$$E[r] = E\left[\sum_{i,j,p}\left[X_i(t) - \left(Y_j(t) + Z_p(t)\right)\right]\right] = 0 \qquad (5)$$

$$E[r] = E\left[\sum_{i,j,p}\left[X_i(t) - \left(Y_j(t) + Z_p(t)\right)\right]\right] = bF\,, \qquad (6)$$

where b is the unknown bias and F is a gross error signature multiplier. The likelihood ratio test statistic is given by equation (7):

$$T = 2\ln\left[\sup \frac{P(r \mid H_1)}{P(r \mid H_0)}\right], \tag{7}$$

which under H_0 follows a chi-square distribution with one degree of freedom. Gross errors are detected whenever T is greater than a predefined critical value, $T_{critical}$. To maintain a given confidence level, α, the test criterion can be set to $\chi^2_{1-\beta,1}$, where $\beta = 1-(1-\alpha)^{1/(r+l)}$ and where r is the number of process units and l is the number of measured variables.

The Tjao-Biegler test is based on the contaminated Gaussian distribution given by equation (8)

$$P(y_i \mid x_i) = (1-\eta)P(y_i \mid x_i, R) + \eta P(y_i \mid x_i, G), \tag{8}$$

where y_i is the measured value of the i^{th} channel, x_i is the true value of the i^{th} channel, $P(y_i \mid x_i, R)$ is the probability distribution function of the random error, $P(y_i \mid x_i, G)$ is the probability distribution function of the gross error, and η is the probability of gross error occurrence. More specifically, the gross error distribution function is given by equation (9)

$$P(y \mid x, G) = \frac{1}{\sqrt{2\pi} b\sigma} e^{\frac{-(y-x)^2}{2b^2\sigma^2}}, \tag{9}$$

where b is the unknown bias. Gross errors are detected whenever the inequality shown in equation (10) is satisfied.

$$\eta P(y_i \mid x_i, G) \geq (1-\eta)P(y_i \mid x_i, R) \tag{10}$$

Maximizing $P(y_i \mid x_i)$ subject to the material or energy balance constraints will produce the gross error estimates.

The neural network test requires the process to be modeled by an artificial neural network. A complete description of neural network modeling methodology is provided in the *OptiRamp* Modeling Submodule white paper. Gross errors are identified whenever errors between fitted and observed process values are outside the norm.

Once all gross errors are eliminated, the process of data reconciliation continues until all measured channels possess trustworthy signals and the material or energy balance is adjusted. The MBR Submodule is able to perform balance adjustments in real time, on demand, and off-line during system setup and integration.

About Statistics & Control, Inc.

S&C—an engineering consulting and technology company headquartered in West Des Moines, IA—solves complex challenges for customers through its unique technology and its highly seasoned team of professionals. The company has a global portfolio spanning the energy, oil and gas, utility, and digital oil field industry sectors. S&C provides clients with turbomachinery control solutions that easily integrate with the existing system as well as *OptiRamp*® solutions, which focus on process and power analytics to optimize processes and, in turn, reduce costs and increase reliability. S&C also provides consulting, dynamic system studies, modeling, automation, training and OTS, and support services.

Statistics & Control, Inc.
4401 Westown Pkwy, Suite 124
West Des Moines, IA 50266 USA
Phone: 1.515.267.8700
Fax: 1.515.267.8701