

Chapter 2

Process Observers and Data Reconciliation Using Mass and Energy Balance Equations

Daniel Hodouin

Abstract This chapter is devoted to data reconciliation for process audit, diagnosis, monitoring, modeling, advanced automatic control, and real-time optimization purposes. The emphasis is put on the constraints of mass and energy conservation, which are used as a foundation for measurement strategy design, measured value upgrading by measurement error filtering techniques, and unmeasured process variables estimation. Since the key variables in a mineral processing unit are usually flowrates and concentrations, their reconciliation with the laws of mass conservation is central to the discussed techniques. Tools are proposed for three different kinds of operating regimes: steady-state, stationary and dynamic. These reconciliation methods are based on the usual least squares and Kalman filtering techniques. Short examples involving grinding, flotation, leaching and thermal processes are presented to illustrate the problems of data reconciliation, sensor placement, fault detection and diagnosis. Strategies for coupling data reconciliation with real-time optimization and automatic control techniques are also proposed. A nomenclature section is included at the end of the chapter

2.1 Introduction

The production goal of a mineral or metallurgical plant (MMP) is ideally to maintain complex unit operating conditions at values where some plant performance index is optimized. The performance index could be expressed either by technical factors, such as the tonnage of valuable material produced, or by the quality of the material produced (*e.g.*, concentrate grade or metal purity). More globally, since a trade-off between the productivity, the material quality, and the production costs is required,

Daniel Hodouin

Département de Génie des Mines, de la Métallurgie et des Matériaux, Université Laval, Québec City, Canada, e-mail: daniel.hodouin@gmn.ulaval.ca

the performance is frequently expressed as an economical index which embeds all these aspects. When the operating conditions drift away from the range of optimal performance, the plant experiences economic losses. The ability of a plant to remain in the vicinity of its optimum operation is related to real-time decision making processes, *i.e.*, to production supervisory systems, real-time optimization systems, and automatic control strategies. Regardless of the strategy used for maintaining a plant close to an optimum performance, the variability around the optimum value relies in the first place upon an efficient evaluation of the performance index—in other words upon the variance of its estimate. The lower the variance, the better is the plant performance.

A plant performance index observer uses the available measurements of the process variables. Typically, in a metallurgical plant, these variables are overall material, phase, and metal flowrates, material chemical compositions, energy flowrates, temperatures, consumed power, *etc.* As any other observer, the plant performance observer simultaneously uses measured values and process models. These models are required to cope with common data processing problems such as measurement uncertainties – which are quite large in a metallurgical operation – lack of measurement availability for critical variables (obviously a performance index is usually not directly measurable), limited knowledge of the process behavior (a difficult problem, particularly in extractive metallurgy), and information redundancy in the available measurements and prior process knowledge. As the process model uncertainties are very large in metallurgical industries, it is common practice to use only constraints, *i.e.*, sub-models – in the sense that they are not causal models as assumed in traditional model-based control and observation. Since the level of confidence in these sub-models must be high to prevent distorting the data set information content by uncertain models, the selected constraints are essentially laws of mass and energy conservation. In the metallurgical, and more generally chemical industries, these observation methods are called reconciliation methods, in the sense that they reconcile the measurement data with the laws of mass and/or energy conservation.

Estimation of process states is required for process performance audit, process modeling, monitoring, supervision, control, and real-time optimization. Whatever the process scale, laboratory, pilot, or full industrial scale, the first step of state estimation is to collect experimental data. Unfortunately, and this is particularly true at the industrial scale, measurements are extremely difficult and inaccurate in the metallurgical engineering field. Production units treat materials that are multi-phase and usually contain extremely heterogeneous particulate phases [1]. The data is highly inaccurate and incomplete, and requires to be improved before being used in the above mentioned applications. The usual statement “Garbage in, garbage out” is particularly true in this context, use of poor data leading invariably to poor models, poor decisions, and improperly designed and operated systems. Therefore, using additional information to the experimental data through process prior knowledge leads to better state estimates. Mathematical models are usually the most efficient way to encapsulate process behavior knowledge. Unfortunately, in metallurgical processes, the knowledge is frequently fuzzy and less accurate than in mechanical and chemical industries.

Figure 2.1 summarizes the concept under discussion in this chapter. The core of a data reconciliation procedure is a mathematical algorithm that can be called either an observer, or an estimator, or a filter. It is an observer in the general sense that it allows the process state observation, *i.e.*, the observation of the variables upon which the process behavior and performance are qualified. It is an estimator in the sense that it estimates numerical values of state variables which may not necessarily be measured or measurable. It is a filter in the sense that, if a state variable is measured, it will correct the experimental value of this process state variable. In this chapter the words observer, estimator and filter, as well as data reconciliation, will be commonly used, without strict meaning differences. Generically, Y is the measurement vector, X the state vector, and \hat{X} its reconciled (or estimated, or filtered, or observed) value. The constraint equations $f(X) = 0$, normally a sub-model as mentioned earlier, are here mainly mass and energy conservation equations. Figure 2.1 presents a steady-state reconciliation (SSR) procedure, but it will be seen later on that stationary and dynamic reconciliation methods can also be considered when the process is not operating in steady-state conditions.

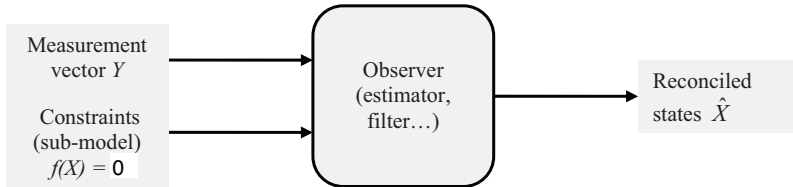


Figure 2.1 Scheme of a data reconciliation procedure using mass and energy conservation constraints

An introductory example. Before going into deeper and more rigorous definitions of the concepts used in data reconciliation, let us give a simple example for qualitatively introducing the key words used in this chapter. The considered plant is the flotation unit of Figure 2.2, and the corresponding data is given in Table 2.1: measured value, measured value standard deviation, as well as reconciled values as explained in Section 2.8.2.

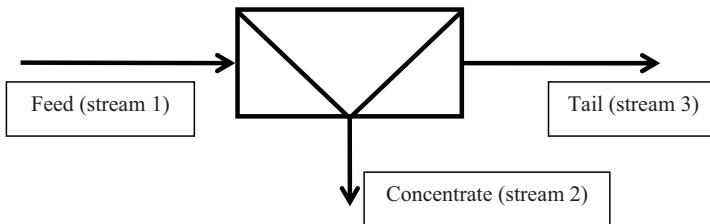


Figure 2.2 A flotation unit

Table 2.1 Measured and reconciled values for the flotation unit of Figure 2.2

Process variable	Flowrates $F(t/h)$			% Cu (100c)			% Zn (100z)		
	Meas. values Y	error s.d.	Reconc. values \hat{X}	Meas. values Y	error s.d.	Reconc. values \hat{X}	Meas. values Y	error s.d.	Reconc. values \hat{X}
Feed	100	1	100.0	2.39	0.05	2.36	8.53	0.19	8.57
Concentrate		0	9.25	23.2	1,11	24.36	7.54	0.34	7.53
Tail		0	90.75	0.12	0,01	0.12	9.65	0.98	8.68

The nine state variables X are the three ore mass flowrates F and the six copper and zinc mass fractions c and z . The measured process variables are a subset of the state variables. The measurement vector Y contains the measured values of the ore feedrate and of the six metal mass fractions. The measurement values are assumed to give an image of the process steady-state behavior. Hence, the constraints of mass conservation $f(X) = 0$ are

$$F_1 - F_2 - F_3 = 0, \quad (2.1)$$

$$F_1 c_1 - F_2 c_2 - F_3 c_3 = 0, \quad (2.2)$$

$$F_1 z_1 - F_2 z_2 - F_3 z_3 = 0. \quad (2.3)$$

Considering this selection of Y and X , the constraints have a bilinear structure. The information content (measurements + constraints) is said to be redundant since it contains two unknown state variables (F_2 and F_3) and three equations. The system is said to be observable since F_2 and F_3 could be estimated by resolving Equations 2.1 and 2.2 or 2.1 and 2.3. The first case leads to

$$F_2 = 9.83t/h ; F_3 = 90.17t/h$$

and the second one to

$$F_2 = 53.1t/h ; F_3 = 47.9t/h.$$

The conflict between these two possible solutions explains why there is a need for reconciling the measurements with the constraint equations. Eliminating unmeasured process states from Equations 2.1 to 2.3 gives

$$c_1 z_3 - c_1 z_2 + c_2 z_1 - c_2 z_3 - c_3 z_1 + c_2 z_2 = 0. \quad (2.4)$$

This equation is called a redundancy equation because it contains only measured quantities. Since there is only one redundancy equation in this case, the redundancy degree of the system is 1. In (2.4), the substitution of the process variables by their measured values gives a value different from zero (2.1×10^{-3}) because of the conflict existing between constraints and measurements. This residual variable generates the parity space of the system. When using reconciled values, Equation 2.4 is exactly verified, as obviously are also the constraints (2.1) to (2.3). The application of a reconciliation procedure to this system generates the following advantages:

- Unmeasured process variables F_2 and F_3 are estimated.
- Process variable measured values are corrected.
- Reconciled values are consistent with the law of mass conservation.
- Plant performance, such as copper recovery, calculated with reconciled data becomes independent of the calculation path (using copper data to estimate the flowrates gives a recovery of 95.4%, and, using zinc data, gives an absurd value of 515%, whereas the reconciled value gives 95.9%). Furthermore results are more accurate than values calculated by any other methods.
- Models that are subsequently estimated from reconciled data (for instance flotation kinetic models) are much more reliable.
- Decisions made from reconciled data are necessarily more efficient than decisions made from raw data.

Reconciliation of mass and energy balance with raw data is a technique which has been known for a long time in mineral processing (see, for instance, [2–4]), but it was already mentioned in chemical engineering as early as 1961 [5]. Crowe wrote a good survey paper in 1996 [6], but the first organized books addressing the chemical process reconciliation topic appeared only at the end of the 20th century [7, 8]. However, the contributions of the MMP community to the field of data reconciliation is mentioned only in the first one. Off-line reconciliation methods for steady-state processes are now quite mature and various computer packages are available. For example, in the MMP field, although they are not at all limited to these applications, one can mention: Bilmat™ and Metallurgical Accountant™¹ (Algosys) [9], Bilco™ and Inventeo™² (Caspeo) [10], JKMultibal™³ (JKTech) [11], Movazen™ (Banisi) [12] and more chemical process oriented: Sigmafine™⁴ (OSIsoft) [13], Datacon (IPS)[14], Advisor™⁵ (AspenTech) [15], and VALI™⁶(Belsim) [16].

The most usual reconciliation techniques are based on the minimization of quadratic criteria, therefore assuming that uncertainties mainly belong to Gaussian distributions. This chapter focuses only on this type of approach. However, alternative reconciliation methods based on artificial neural networks have also been proposed. But those do not offer either the same rigorous statistical and physical background or the same result reliability analytical evaluation tools as in the approach presented here (see, for instance, [17–20]). Linear matrix inequality (LMI) methods have also been proposed by Mandel *et al.* [21].

Steady-state methods are applied off-line to mineral processes such as comminution [22, 23], flotation [59], gold extraction [25–28], hydrometallurgy [29] and [30], pyrometallurgy [31–33], and cement preparation [34]. On-line applications to steady-state processes are actively used, while stationary-state methods that make a

¹ Bilmat and Metallurgical Accountant are registered trademarks of Algosys, www.algosys.com

² Bilco and Inventeo are registered trademarks of Caspeo, www.caspeo.net

³ JKMultibal is a registered trademark of JKTech, www.jktech.com.au

⁴ Sigmafine is a registered trademark of OSIsoft, www.osisoft.com

⁵ Advisor is a registered trademark of AspenTech, www.aspentech.com

⁶ VALI is a registered trademark of Belsim, 174k rue De Bruxelles, 4340 Awans, Belgium, www.belsim.com

trade-off between model uncertainties and measurement errors are well developed but still not frequently used [24, 35]. Dynamic methods have been proposed and are being developed [36–40], but applications of those are not documented.

Other topics are related to data reconciliation, such as sampling error and reconciliation criterion weighting factor evaluation [41, 79, 96], reconciled value accuracy evaluation [44], use of reconciled values to calculate and display plant performance indices, such as concentrate grade and recovery. Owing to their better reliability, these indices may improve manual or automatic process performance optimization [45, 46]. Coupling of dynamic reconciliation with control has also been investigated [47, 48], as well as gross error detection, fault isolation and diagnosis [49, 50]. Finally, instrumentation design can be performed on the basis of data reconciliation methods used as process observers [51, 52].

The chapter includes the following parts. Section 2.2 begins with definitions of plant process variables and operating regimes that may be considered in reconciliation methods. Then, in Section 2.3, mass and energy conservation equations are written for different plant operating modes. As is the case for all the sections of this chapter, strong emphasis is placed on mass balance problems, rather than on energy. Then, since the basic incentive for reconciling data is the presence of measurement errors, Section 2.4 covers measurement problems, while Section 2.5 presents the observation equations. Section 2.6 introduces the general principles of data reconciliation algorithms based on least-squares procedures, while Section 2.7 gives the steady-state and stationary operating regime solutions for the linear constraints case. Section 2.8 briefly discusses the non-linear reconciliation cases. Section 2.9 is devoted to the reliability of reconciled data analysis. Section 2.10 presents some reconciliation methods for plants operating in the dynamic regime, while Section 2.11 briefly addresses the issue of how reconciliation methods can help to improve metallurgical plant instrumentation design strategies. Section 2.12 explores how mass and energy conservation constraints can also be used for detecting abnormal process behaviors or measurement problems. Finally Section 2.13 makes way for the integration of reconciliation techniques into optimization or control loops.

2.2 Process Variables and Operating Regimes

Variables involved in mineral processing units characterize process states relatively to material quantities (extensive properties) or qualities (intensive properties). Mass or volume flowrates and hold-ups of solid, liquid, slurry and gas phases (or of given species within these phases) belong to the first category, as well as other related variables such as levels and flowrates. The second category frequently includes the following variables:

- concentrations of chemical species or minerals in solid, liquid and gaseous phases;
- solid percentage in slurries;
- particle size and density distributions;

- concentrations of chemical species or minerals in particle size or density classes;
- pressures, temperatures.

Process variables may or may not be directly measured by sensors or analytical devices. For instance, ore flowrates may be obtained through the information given by a slurry volume flowmeter, a density gauge, and ore specific gravity measurement. Metal flowrates are other examples of process variables that are obtained through a combination of measurements (flowrates and concentrations).

There is no unique way to select the set of variables X that characterize process states. The structure of the models and constraints describing a process behavior depends upon variable selection. This will have an impact on the observation or data reconciliation method, although the resulting values of reconciled states should not rest upon the problem formulation, if consistent information processing methods are used. Similarly, the measured values Y of the Z process variables used as input to the reconciliation procedure may or may not be raw measurements of the process states X . Furthermore, they can be obtained by combining several sources of raw measurements. The structure of the database Y and of its uncertainties may have a significant impact on the reconciliation method and sometimes on the reconciled results.

The variation of the process states X as a function of time depends on the intrinsic dynamics of the process, on the variations of the operating conditions applied to the process, and on the disturbance dynamics. The process operating regimes can be classified into six types:

- the steady-state regime, when all the process input and state variables are constant;
- the stationary regime, when the process dynamics are limited to random variations around a steady-state regime;
- the transient regime, where the process evolves from one steady-state to another one;
- the quasi-stationary regime, which corresponds to stationary random variations around persistent mean value changes;
- the cyclic regime, when the process operates cycles of production, such as in the carbon-in-pulp process where the carbon transfer is cyclic, or in smelting processes where the material is cyclically cast;
- the batch regime, when the states evolved according to a trajectory from an initial state to a final state.

Figure 2.3 illustrates process state variations for four different operating regimes. Local stochastic state variations are mainly due to input disturbances (ore grade variations for instance), while the trends are mainly the results of deterministic changes of the manipulated variables. Since in real processes it is impossible to maintain strictly constant conditions, the steady-state regime corresponds to virtual operating conditions.

The statistical properties of any stationary process variable, a scalar or a vector x (either input disturbances and/or process states and/or process outputs), can be

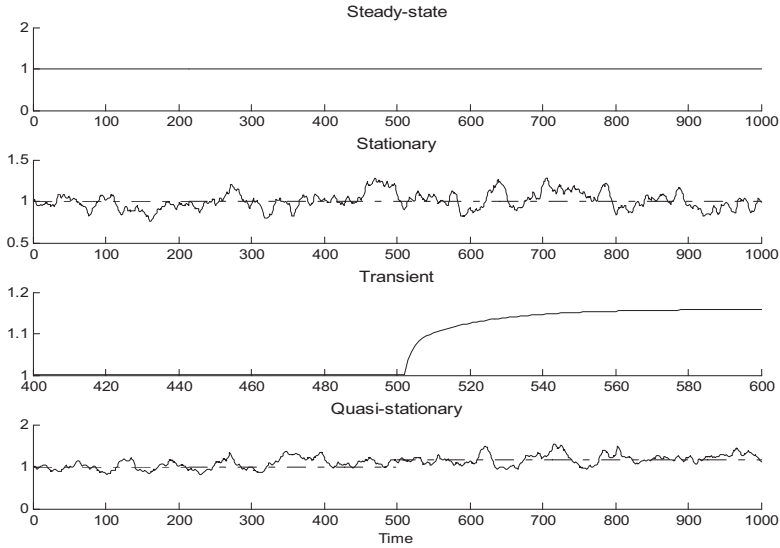


Figure 2.3 Typical variations of process variables for various operating conditions

described by the usual statistical features: the mean μ_x , the variance $V_x(0)$ and the autocovariance $V_x(k)$ (or autocorrelation $\rho(k)$ if normalized by the variance):

$$\begin{aligned} E(x) &= \mu_x; \\ E((x(t) - \mu_x)(x(t-k) - \mu_x)^T) &= V_x(k). \end{aligned} \quad (2.5)$$

The nature of the reconciliation procedure to be applied to filter industrial data must be adapted to the operating regime that was prevailing during data gathering, and to the measurement strategy then applied. There is no systematic method to decide whether a steady-state, stationary or dynamic filter must be applied, but some hints can be helpful. Some of the conditions for the application of a steady-state observer are:

- The process deviation from a theoretical steady-state is of low magnitude compared with measurement error amplitudes (variance matrix V). In other words, the diagonal terms of V are large in comparison with the diagonal terms of $V_x(0)$.
- The process deviation from a theoretical steady-state is significant with respect to the measurement errors, but the dynamic variations are produced by stationary disturbances of high frequency spectra in comparison with the natural process dynamics – in other words, the process variable autocorrelogram time widths are small in comparison with the width of the cross-correlation between process inputs and outputs. However, instantaneous measurements data set could be processed by the steady-state method providing that the high frequency disturbance variances are added to the measurement error variances.

- The process is operating in a stationary regime while the frequency of the variations around the steady-state is of the same range as the process natural dynamics and of significant amplitude. The steady-state method could be applied to averaged measurements in a time window sufficiently large to significantly dampen the dynamics of the variations around the steady-state values. The time window must be at least two times larger than the process time constants. Measurement variances should be augmented to take account of the residual dynamics of the variations around the steady-state values.

Obviously, stationary methods could be applied to instantaneous measurements when the process is stationary, *i.e.*, when the process is not operating during significant changes of the underlying steady-state values. In other words, the process must be sampled at a time sufficiently far from a deterministic change of the average operating conditions. When the process is clearly in a transient regime or cyclic or batch, a dynamic reconciliation method should be used to take account of the lags occurring between the various process states.

2.3 Models and Constraints

The component stream networks, where the material components (or enthalpy) are flowing through the industrial unit, are described as oriented graphs, made of p branches representing streams, and n_n nodes representing accumulation equipments. In the following, emphasis is placed on mass balance to alleviate the presentation. Energy balance equations have the same structure as mass balance equations, enthalpy (or heat) being added to the list of the $n + 1$ components (including total mass) that must be conserved. Since reconciliation techniques are quite similar when energy balances are considered, there is no need to repeat the expression “mass and/or energy balance” throughout the text. The selected components used for writing balance equations can be either phases, or species as minerals, metals, atoms, molecules, ions, or classes of physical properties (size, density), and enthalpy. Each component may have its own network characterized by an incidence matrix M_i , whose entries 1, -1 , and 0 represent either a node input, or output, or a not connected stream. The process states used to write the mass conservation constraints are usually mass flowrates and mass fractions, therefore leading to bilinear equations. When considering component flowrates – instead of total flowrates and component mass fractions – the state equations can be kept linear. This case will be considered first.

2.3.1 The Dynamic Linear Mass Balance Equation

The dynamic mass conservation equations for any i component (including phases, species, properties and the total material flowing in the various streams) are

$$\frac{dm_i}{dt} = M_i f_i + P_i - \varepsilon_i, \text{ for } i = 0 \text{ to } n + 1, \quad (2.6)$$

where m_i is the component i mass vector (including the total mass) accumulated at the network nodes, f_i the stream mass flowrate vector, P_i the component i production rate vector at the various nodes (for the total material $P_0 = 0$), and finally ε_i an uncertainty vector providing for structural errors such as forgotten secondary streams or intermittent streams, or errors in the production rates evaluation.

Usually rates of production are unknown and cannot be measured independently from other states. There are three possible situations:

1. the component i is transformed, and the conservation equation must not be written at the corresponding node;
2. the transformation is of very low magnitude and is simply incorporated into ε_i :

$$\frac{dm_i}{dt} = M_i f_i - \varepsilon_i, \text{ for } i = 0 \text{ to } n; \quad (2.7)$$

3. the component i is not transformed at a given node, and $P_i = 0$. Therefore, in the absence of structural uncertainties, this leads to the exact dynamic conservation constraint:

$$\frac{dm_i}{dt} = M_i f_i, \text{ for } i = 0 \text{ to } n. \quad (2.8)$$

Defining the state vector as $x_i = (m_i^T, f_i^T)^T$, Equation 2.8 can be written in the generic form:

$$E_i \frac{dx_i}{dt} = D_i x_i, \text{ for } i = 0 \text{ to } n + 1. \quad (2.9)$$

Rather than a model allowing process simulation, this is a singular model, *i.e.*, a set of constraints linking the state variables. Its discrete version is

$$E_i x_i(t + 1) = F_i x_i(t), \text{ for } i = 0 \text{ to } n + 1. \quad (2.10)$$

Example. In a complex ore comminution or separation plant, the conservation constraints could, for instance, be written for the following components: slurry, water, ore, copper, lead, zinc, gold, particle size classes, and gold in particle size classes. If ten size classes are defined, the number of component conservation equations is 27 ($n + 1 = 27$). As will be discussed in Section 2.3.5, the component definition selected here will create additional constraints since, among others, the gold species is selected at two different levels of the mass balance equations.

2.3.2 The Linear Stationary and Steady-state Cases

When the process is in a stationary operating regime, *i.e.*, a regime randomly fluctuating around a steady-state, the rate of accumulation dm_i/dt can be omitted and incorporated into the uncertainties ε_i :

$$M_i f_i = \varepsilon_i, \text{ for } i = 0 \text{ to } n + 1, \quad (2.11)$$

where ε_i is interpreted as a stochastic term containing the neglected accumulation rate, production rate and structural uncertainties. When the process is strictly operating in a stationary mode, as defined in Section 2.2, ε_i can be interpreted as a stationary random signal with the following statistical properties deduced from the statistical variations of the component flowrates that are usually assumed to behave as normal random variables. If V_f is the covariance matrix of f , the stacked f_i vectors, then the stacked vector ε of the ε_i s has the following properties:

$$\varepsilon \sim N(0, V_\varepsilon), \text{ with } V_\varepsilon = M V_f M^T, \quad (2.12)$$

where M is the block diagonal matrix of the M_i s.

The steady-state case is then a particular case of the stationary equations, when V_ε has a zero value.

2.3.3 The Bilinear Case

Instead of using, as state variables, the component flowrates (including the total mass flowrate), one can use the total mass flowrates and the phase or species mass fractions. Obviously the models are strictly identical, but the selection of these more usual variables changes the structure of the equations with respect to the state variables. For the total mass conservation, the equations are now

$$\frac{dm_0}{dt} = M_0 f_0 - \varepsilon_0, \quad (2.13)$$

where M_0 is the total mass incidence matrix. For the phases or species mass fractions c_i in the streams, and h_i in the node loads, the equations are

$$\frac{dm_0 \bullet h_i}{dt} = M_i(f_0 \bullet c_i) + P_i - \varepsilon_i, \text{ for } i = 0 \text{ to } n, \quad (2.14)$$

where \bullet is Hadamard's product. One can also incorporate Equation 2.13 into (2.14) which becomes

$$m_0 \bullet \frac{dh_i}{dt} = M_i(f_0 \bullet c_i) - h_i \bullet (M_0 f_0 - \varepsilon_0) + P_i - \varepsilon_i \text{ for } i = 0 \text{ to } n. \quad (2.15)$$

Dynamic equations similar to (2.7) and (2.8) can also be obtained. More particularly the stationary case becomes

$$M_0 f_0 = \varepsilon_0, \quad (2.16)$$

$$M_i(f_0 \bullet c_i) = \varepsilon_i, \text{ for } i = 1 \text{ to } n, \quad (2.17)$$

where it is clear that these equations contain flowrates and mass fractions cross-products, thus giving a bilinear to the equation set.

Example. For the same example as in Section 2.3.1, f_0 would be the ore flowrate, and c_i would correspond to copper, lead, zinc, gold and the 10 size class mass fractions. To keep the bilinear structure, it is necessary, in this particular case, to ignore the data levels corresponding to slurry and gold content in size classes.

2.3.4 Multi-linear Constraints

Process measurements usually concern flowrates and material compositions. In that case, process states are defined with two levels of properties as depicted in the bilinear case above. Unfortunately, when the performances of a mineral or metallurgical processing plant must be deeply assessed, more than two levels of material properties need to be handled [53]. Material streams may contain various phases (ore, carbon, aqueous, organic, and gas phases) which are characterized by many properties such as flowrates, particle density and size distributions, as well as mineral and chemical compositions of each phase and particle class. The mass conservation equations complexity can rapidly increase with the level of detail needed for process analysis.

Example. Gold ore processing plants involve different phases (slurry, water, carbon and ore) that are characterized by various properties (size distributions, along with chemical and mineral compositions), as well as mineral and chemical compositions of ore and carbon size classes. Only a few studies on data reconciliation in the gold ore processing industry are available [25–27],[54]. Figure 2.4 shows a possible multi-level representation of the stream materials. First, the slurry phase is divided into liquid, ore and carbon phases. Then, each phase is subdivided into its specific components:

- the liquid phase into chemical reagents (CN^- , O_2) and leached species (Au, Ag, Cu);
- the ore phase into populations of particles (such as coarse and fine) which are subsequently split into classes of particles (such as $-38\mu\text{m}$, $+38/-53\mu\text{m}$, $+53/-75\mu\text{m}$), each particle size class being characterized by its mineral (native gold, pyrite, hematite) content, and, subsequently, the mineral metal contents (Au, Ag, Cu);
- the carbon phase into size classes, each class being characterized by its metal content.

In such a complex system with six different levels of information, the mass conservation system becomes 6-linear. In addition, various components are simultaneously considered at various levels, thus creating a complex set of additional constraints that are necessary to ensure gold conservation is consistent between the different levels. This case study is discussed in [28], and the next section gives some more information on the additional constraints that are usual in MMP plants.

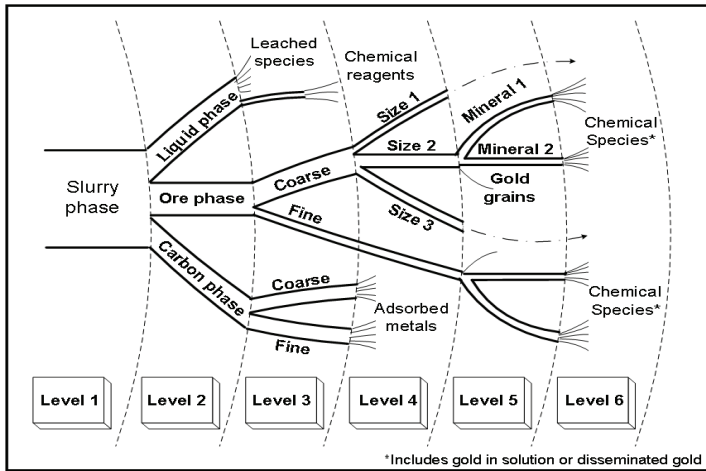


Figure 2.4 Multi-level representation of stream matter for a generic gold ore processing plant

2.3.5 Additional Constraints

In addition to the above basic mass conservation constraints, one may have to add equations that represent either additional structural constraints, or stoichiometric constraints, or process behavior modeling assumptions (mass or energy transfer and rate equations). In the first case, the most frequent constraint arises when an exhaustive species analysis is available. Then the species mass fractions must sum to 1,

$$\sum_i c_i = 1 \quad (2.18)$$

for i values corresponding to exhaustive analyses. Alternatively, the constraint may be expressed by writing that the sum of some component flowrates must be equal to the total flowrate:

$$f_0 = \sum_i f_i. \quad (2.19)$$

In the above example, these constraints are valid for carbon, water and ore components, and also for particle size classes.

A stoichiometric constraint may arise for instance when minerals are assumed to have fixed composition, such as FeS_2 or CuFeS_2 . When both sulfides are present and are the only sulfur, copper, and iron carriers, there is a relationship between sulfur, copper, and iron contents that must be verified. Another possible constraint arises when there are stream splitting systems where an equal share of the streams is assumed. Another possibility is that a species production rate has a prior estimated value that must be obeyed with some level of uncertainty.

Other constraints may arise in multi-phase systems, when a species transfers from one phase to another one. This is obviously the case in leaching, elution, ad-

sorption and solvent extraction processes when a metal transfers from a solid to a liquid phase or from a liquid phase to another one.

When the same species is analyzed at two different levels, for instance in ore and particle size classes, the species content in the ore must be simultaneously consistent with particle size distribution and species assays in size classes. The example of Section 2.3.1 shows a case study for gold, while Hodouin and Vaz Coelho [53] show another example for processing of a uranium ore. In the example of Section 2.3.4 there are many such additional constraints.

Also, when chemical equilibrium between phases is assumed, the thermodynamic equilibrium conditions must be satisfied [29]. Finally, inequality constraints may also be in force. For instance, species mass fractions must have values between 0 and 1, and flowrates must be positive.

2.3.6 Summary of Stationary Conservation Equations

Equations 2.11 or 2.16 and 2.17 coupled to specific additional constraints can be gathered into the following form:

$$f(X) = \varepsilon, \quad (2.20)$$

where X is a vector which gathers all the n_X state variables and f the q conservation equations. In the linear case all the equations are gathered into

$$MX = \varepsilon, \quad (2.21)$$

where M is a matrix of coefficients containing blocks such as M_0 and M_i and possibly other terms depending upon the additional constraints. It may happen that constant terms are also present in the constraints; Equation 2.21 would then become

$$MX = K + \varepsilon. \quad (2.22)$$

To simplify the presentation, this case will not be covered in the following developments. If this is required, the modification of the formulae to take account of K would not be complex.

2.4 Sensors, Measurement Errors and Observation Equations

Measurement errors play a central role in data reconciliation, since one important feature of the technique is to correct data that is contaminated by measurement inaccuracies. This section proposes a description of the statistical properties of measurement errors, with some emphasis on particulate material sampling and analysis.

2.4.1 Statistical Properties of Measurements and Measurement Errors

The measurement error statistical properties, as well as the methods used to process them, depend upon the category they belong to. Measurements can be classified into the following categories:

- On-line or off-line measurement. For instance, flowmeters or density gauges are on-line while chemical or physical analysis of a sample in a laboratory is off-line. Automatic sampling followed by centralized X-ray fluorescence analysis is considered an on-line analysis.
- Measurement either on part of the material, for instance, chemical analysis of an ore sample, or on the whole material, as performed by a flowmeter installed around a pipe.
- Continuous or discrete measurement. A flowmeter or a particle size analyzer delivers a continuous signal, while the analysis of a sample of material taken at constant time intervals is delivered at a given frequency.
- Averaged measurement value or instantaneous measurement. The reading of a flowmeter can be averaged in a given time window, or a sample made of composite increments can be analyzed for its physical or chemical properties. On the contrary a sampled flowmeter signal gives an instantaneous value.
- For discrete measurements, averaged or instantaneous samples can be taken at constant time periods (systematic sampling), or randomly (random sampling), or randomly within constant time intervals (stratified sampling).

Moreover, one can distinguish three main types of measurement errors that must be processed differently by the process observers:

- Systematic errors, or biases, are the consequences of sensor calibration drifts, interaction effects – such as the interaction of foreign chemical elements on the analyzed species – or biases in sampling procedures.
- Random centered errors. They results from many independent sources of noise due to the heterogeneous nature of the material to be analyzed and to the inherent fluctuations of the analytical devices. They have a zero mean value and are usually considered as obeying normal distributions, unless the error variance is large compared with the nominal value of the process variable, in which case the normal distribution might mean that a process variable could have negative values, an unacceptable property for inherently positive variables such as concentrations or flowrates.
- Accidental gross errors, due, for instance, to contamination of samples, tagging mistakes, transmission faults.

For correct use of measurements, systematic errors must be detected, corrected, and their sources eliminated, for instance by maintenance and calibration of sensors, or by redesigning sampling procedures. Accidental gross errors must also be detected, using fault detection and isolation (FDI) techniques, and the corresponding

data eliminated or replaced by interpolation techniques. Random errors are filtered by observers.

The quality of a measured value is characterized by its accuracy, *i.e.*, the absence of bias, and by its reliability, quantified by random error variances. In the following, the statistical properties of any measurement error e will be defined by

$$e \sim N(b, v), \quad (2.23)$$

where N stands for the normal law, b for the bias and v for the variance. For multiple unbiased measurements (vector e), the statistical properties are

$$e \sim N(0, V), \quad (2.24)$$

where, most frequently, V is a diagonal matrix. However, it may happen that the measurement errors of different process variables are correlated. In these conditions V contains non-zero off-diagonal terms. This may arise when a common sample or a common measuring device is used to measure different process variables. The following situations are examples of such correlated errors:

- Particle size analysis by sieving. Particles that are not on the right sieve are necessarily present on another one, thus creating a negative correlation between the errors of fractions belonging to different size intervals [55].
- Measurement systems involving a common sensor, such as an X-ray fluorescence analyzer used at different sampling locations, create correlation between the errors.
- Matrix effects in X-ray fluorescence analysis might correlate concentration measurement errors of different metals.
- Synchronous incremental sampling of different streams may induce error correlation due to intercorrelation of the streams dynamics, which is created by the process itself (see integration error in the following section).

It is important to point out common misinterpretation of error correlation. Even if two process variables are correlated, *e.g.*, concentrations of two different metals in an ore, their measurement errors are not necessarily correlated. If the sampling and analysis steps are uncorrelated, the measurement errors are usually uncorrelated, even when the process variables are correlated.

2.4.2 Measurement Errors for Particulate Materials

Since they are made of randomly organized grains of various minerals, ores are heterogeneous materials. Also, as these grains take random shapes and sizes, local properties of ores may not be representative of their overall average properties. Breakage of ores into smaller particles may help ore homogenization. However, if the particles are not perfectly mixed, local properties of a batch of particles may not

be representative of the average whole batch properties. Based on Gy [56] and Pitard [57], the main error sources for particulate materials are now briefly presented.

Fundamental error. Unless the material is perfectly homogeneous at a micro-scale level, sampling a limited number of particles to characterize the whole properties of an ore batch inherently leads to measurement errors. According to Gy, the errors induced in a sample by ore constitution heterogeneity will be called a fundamental error. It cannot be eliminated. However, it can be evaluated. It has a zero mean and a variance that can be estimated by

$$\sigma_F^2 = E \left[\left(\frac{x_s - x^*}{x^*} \right)^2 \right] = \frac{1}{M_s} c \sqrt{\frac{d_{lib}}{d}} f g d^3, \quad (2.25)$$

where x_s and x^* are the mineral contents, respectively, of the ore sample and the ore batch, M_s the sample mass, d_{lib} the liberation particle size, d the sieve opening retaining 25% of the particles, f the shape factor, g the size distribution factor, and c the composition factor defined as

$$c = ((1 - x^*)/x^*)((1 - x^*)\rho_{min} + x^*\rho_{gan}), \quad (2.26)$$

where ρ denotes the mineral or gangue density.

Integration error. This error is induced by the particle distribution heterogeneity, either spatially distributed for a fixed batch of ore, or as a function of time for a flowing material. Hybrid integration errors may also occur when a sample is taken from a given location of a flowing stream, rather than from all the possible locations in the stream (using, for instance, a cross-stream cutter). This error can be evaluated if the signal heterogeneity statistics, such as mineral content autocovariance, or geostatistical variograms of the particle batch, are known. For a time related concentration of a stream sampled in a time window of width T , the definition of the integration error is simply

$$e_I = \left\{ \begin{array}{l} \text{composition} \\ \text{of the incre-} \\ \text{ment reunion} \end{array} \right\} - \left\{ \begin{array}{l} \text{average stream} \\ \text{composition over} \\ \text{time window T} \end{array} \right\}. \quad (2.27)$$

The variance of e_I can be calculated from the sampling strategy parameters and the autocovariance of e_I . This calculation can be extended to a composition vector of several streams [79].

Materialization error. This arises when extracting the sample from the ore batch to be characterized, *i.e.*, when executing the designed sampling scheme. For example, when cutting a stream flow with a moving sampler, errors may occur if all the particles do not have the same probability of entering into the sampler, either because the sampler opening is too tight or its speed too high or not uniform. Other materialization errors may arise if the sampler is overflowing or when gathering the various sampling increments to generate a composite sample. Spillage and contamination are also materialization error sources. Careful design of sampling devices

and sampling execution allows elimination of this type of error, which is usually almost impossible to quantify.

Preparation errors. They are related to the operations that condition the final sample to be analyzed by the measuring equipment (X-ray fluorescence device for instance). The preparation involves secondary sampling, and thus induces new fundamental and integration errors. Drying, grinding are also operations that may induce contamination or transformation of the phases, by oxidation for example.

Analysis error. Finally, as the analytical method is based either on mechanical methods (sieving for instance) or chemical principles (titration) or spectral analysis (X-ray fluorescence, atomic adsorption), the analyzing device itself necessarily induces errors. Usually, these errors can be quantified but cannot be avoided.

Total measurement error. The variance of the resulting measurement error is obtained by summing the variances of all the contributing errors.

2.5 Observation Equations

The general form of a process observation equation is

$$\begin{aligned} Z &= g(X), \\ Y &= g(X) + e, \end{aligned} \tag{2.28}$$

where X is the vector of state variables, Z the measured process variables and Y the measurement value of Z . There are two different ways to define the measurements. Logically, the measured variables should be selected as the process variables that are directly measured by on-line sensors or laboratory analytical instruments. However, there is frequently some raw measurement preliminary processing. For instance the particle masses retained on sieves are converted to mass fractions [58], or slurry volume flowrates and densities converted to ore flowrates, thus assuming that the ore specific mass is a known parameter. The propagation of the measurement errors through the preliminary processing must be evaluated properly, since the calculation process may not only increase the variance, but also create covariance terms in the V matrix. These covariance terms structure the reconciliation results and should not be ignored, as shown for instance by Hodouin *et al.* [103] and Bazin and Hodouin [55] for particle size distributions.

Moreover, when the measured variables are not state variables, one may find it easier to combine measured variables in such a way to obtain state variable measurements even if those have not been directly measured. An advantage of this procedure is that the observation equation becomes linear and can be written as

$$\begin{aligned} Z &= CX, \\ Y &= CX + e, \end{aligned} \tag{2.29}$$

where C is a matrix of coefficients with values pointing at measured state variables. A possible drawback of the method is that the covariance of the pseudo-measured

states should be calculated from the variance of the measurement errors of the source measured variables.

Alternatively, as the state variable selection is not unique, one may also select the state variables in such a way that they are confounded with directly measured variables. The drawback of this procedure is that it usually does not allow state equations to be linear. In summary, selection of the X and Y sets is not unique, but it must be made such that the functions f and g and the X and Y uncertainties covariance matrices be the least complex possible. Obviously, there is no perfect selection, since simplifying one function necessarily implies that the other one becomes more complex.

2.6 General Principles of Stationary and Steady-state Data Reconciliation Methods

The core of a data reconciliation procedure is a model-based observer that makes use simultaneously of process models (stationary regime constraints as defined in Section 2.3.6) and measurements as defined in Section 2.5. It optimally estimates unmeasured process variables in such a way that the data is reconciled with the process model, while respecting the measurement and model uncertainties (see Figure 2.5). The observer is based on the minimization of a reconciliation criterion, which usually consists of a quadratic sum of residuals $J(X)$ containing both the node imbalances ε and the measurement errors e . The reconciliation problem is formulated as

$$\begin{aligned}\hat{X} &= \arg \min_X [(Y - Z)^T V^{-1} (Y - Z) + \varepsilon^T V_\varepsilon^{-1} \varepsilon] \\ Z &= g(X) \\ Y &= Z + e; \quad e \sim N(0, V) \\ \varepsilon &= f(X); \quad \varepsilon \sim N(0, V_\varepsilon)\end{aligned}\tag{2.30}$$

subject to

$$X_{min} \leq X \leq X_{max},\tag{2.31}$$

where X are the plant states, Y the measured values of Z , e the measurement errors assumed to have zero mean values and known variance matrix V , and ε the constraint uncertainty values assumed also to have zero mean values and known variance matrix V_ε . State equation $f(\cdot)$ consists of mass conservation constraints and additional constraints (as well as energy constraints when needed), while observation equation $g(\cdot)$ relates measured variables to state variables. Finally, estimated plant states X have to be within physically meaningful intervals. For instance, mass fractions have to be between 0 and 1 and flowrates should have positive values.

The criterion in Equation 2.30 can be viewed as an empirical least-squares procedure or as the maximum likelihood solution of the state estimation problem if measurement errors and model uncertainties are Gaussian. The latter might be not strictly verified since model structure, parameters, and neglected dynamics may not be Gaussian, as well as measurement error uncertainties, which, obviously, can-

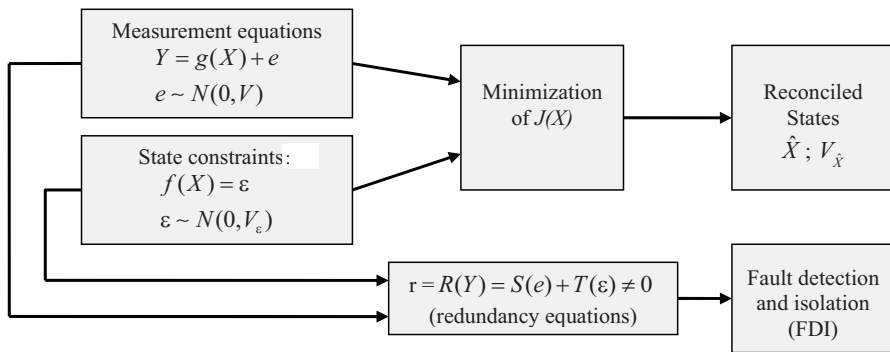


Figure 2.5 General scheme for stationary data reconciliation and fault detection and isolation (FDI)

not be anything between minus and plus infinity, because of physical inequality constraints. Without pretending that this is a statistically correct statement, it is in practice observed that this criterion is sufficiently powerful for improving data at industrial or even lab scales. Alternative techniques such as LMI [21], robust estimators [60], and artificial neural networks have also been proposed [17–20],[61].

In parallel with the reconciliation procedure, Figure 2.5 also shows optional FDI procedures. Redundancy equations R are obtained by elimination of X between the measurement equations and state constraints. Because of the conflict generated by the uncertainties e and ε , these equations generate residual values that are not zero but functions of e and ε . These residuals can be used to detect measurement biases or abnormal deviations to mass and energy conservation laws. These concepts will be discussed later on in Section 2.12. In Figure 2.5, the term $V_{\hat{X}}$ is the variance matrix of the reconciled values that will be discussed in Section 2.9.

The stationary optimization problem defined by Equation 2.30 degenerates into two limit reconciliation problems when e or ε are assumed to have null values: the steady-state case and the node imbalance case.

Steady-state data reconciliation. The SSR case is obtained by setting V_{ε} to zero, thus removing the second term of the criterion. As already said the steady-state case is an ideal situation. There is a continuum between the stationary and the steady-state case, and one can superficially say that SSR is legitimate when V_{ε} is small in comparison with V .

Node imbalance data reconciliation. At the other end of the relative values of V and V_{ε} spectra, one may consider the case where the measurement values are much more accurate than the conservation constraints. In other words V is small in comparison with V_{ε} . The first term of the reconciliation criterion disappears and only the residuals ε are estimated. These estimates are called the reconciled node imbalances.

The feasibility of reconciliation procedures is governed by process observability [62, 63] and, its corollary, information redundancy. These two concepts are now discussed briefly.

2.6.1 Observability and Redundancy

A process state x_i is said to be observable or estimable if it can be calculated using simultaneously the measurement values and the conservation constraints, or part of them. Therefore the problem is to find a unique estimate of the variables x_i of the vector X that satisfies the following system:

$$\begin{aligned} f(X) &= 0, \\ g(X) &= Z, \end{aligned} \quad (2.32)$$

where Z is known and is the exact value of the measured variables. Equation 2.32 represents the constraint and observation equations of (2.30) where the uncertainty variables are set to their most probable values, *i.e.*, zero. A state variable, such as a metal concentration, which is directly measured, is obviously observable since one possible estimate is its measured value. Hence, the concept of observability is only important for state variables that are not directly measured.

When, in system (2.32), there is at least one of the equations (state or measurement equation) that cannot be removed without losing x_i observability, x_i is said to be non-redundant. When there is more than one possible way to estimate the value of a state, using different equations of the system (2.32), this state variable is said to be redundant. When the state is directly measured, it is redundant when it is still possible to estimate its value in the case the measurement is unavailable. Because of the inherent uncertainties of the Y values, the estimate value obtained from the following system for a redundant state x_i :

$$\begin{aligned} f(X) &= 0, \\ g(X) &= Y, \end{aligned} \quad (2.33)$$

depends on the subset of equations that is kept for calculating the state variable.

When all the state variables are estimable, the process is said to be observable. When at least one state variable is redundant the process information is said to be redundant. When all the state variables are observable and non-redundant, the system is said to be of minimal observability.

One can define the process information overall degree of redundancy as the largest number of equations that can be eliminated from the system without losing process observability. Usually it is related to the number of equations minus the minimum number of required equations to obtain minimal observability. Redundancy degrees for individual states can also be defined [64].

The redundancy degree is strongly coupled to the data reconciliation performance: the higher the redundancy, the higher the reconciled value reliability (for state estimate reliability, see Section 2.9). Moreover, the higher the redundancy, the higher the robustness of the observer. This means roughly that the number of possible sensor failures that do not hinder process observability increases with the degree of redundancy [64]. Assuming that all the variables are observable, a redundancy degree can be defined as

$$\rho_r = \frac{q + m - n_X}{q}, \quad (2.34)$$

where q and m are, respectively, the number of state equations and observation equations, and n_X the number of state variables. This index varies between 0 and 1 from the overall minimal observability to the case where all the states are measured directly or indirectly. The determination of an unmeasured state observability may be a difficult task for non-linear f and g functions. Let's have a look at the linear case.

The linear case. The system of equations containing the plant information is

$$\begin{aligned} MX &= 0, \\ CX &= Z, \end{aligned} \quad (2.35)$$

It can be globally rewritten as

$$\Psi X = \Phi Z. \quad (2.36)$$

A state variable x_i is observable if there is at least a subset of equations in (2.36) that allows the calculation of x_i when Z is known. The state vector would be observable if the rank of the matrix Ψ is n_X , the number of state variables. If the process is globally at minimal observability, then Ψ is an invertible matrix (regular matrix).

Redundancy equations. Another way of looking at redundancy is to eliminate the state variable X from the system (2.32) (or (2.35)) in the linear case). The remaining set of equations is

$$R(Z_r) = 0 \text{ or } RZ_r = 0 \text{ in the linear case.} \quad (2.37)$$

$R(\cdot)$ is the set of redundant equations, and Z_r is here the generic vector of the measured process variables which are redundant (Z or a subset of Z). The number of equations it contains is usually $m + q - n_X$, where $n_X = (n + 1) \times p$ in the linear case with $(n + 1)$ components and p streams. The number of redundant equations is thus directly related to the redundancy degree of the reconciliation problem. When replacing Z_r by the measurement values Y_r , the system (2.37) is no longer verified because of the unavoidable measurement errors. The resulting vector is a residual, that is a vector of a space called parity space, physically related to nodes or joint nodes imbalances. It is a function of the measurement errors and model uncertainties insofar as stationary conditions are assumed. It vanishes when the uncertainties have zero values:

$$R(Y_r) = r = S(e_r) + T(\varepsilon) \text{ or } RY_r = Re_r + T\varepsilon \text{ in the linear case.} \quad (2.38)$$

The linear stationary case with $Z = X_m$. Let us consider, as an illustrative case, the situation where the measured variables are state variables (X_m), and the stationary conservation constraints are linear, and gathered into

$$MX = \varepsilon. \quad (2.39)$$

The vector X can be reorganized as

$$X = \begin{pmatrix} X_m \\ X_{um} \end{pmatrix}, \quad (2.40)$$

where X_m is the set of measured variables, *i.e.*, the vector Z of this particular case, and X_{um} the set of unmeasured variables. The matrix M can also be decomposed into two sub-matrices, one for X_m , and the other one for X_{um} :

$$M = (M_m, M_{um}). \quad (2.41)$$

This allows Equation 2.39 to be written as

$$M_m X_m + M_{um} X_{um} = \varepsilon. \quad (2.42)$$

Providing that the system is at minimal observability, *i.e.*, that the measured values are properly placed on the state vector, M_{um} is invertible and X_{um} calculable as

$$X_{um} = (M_{um})^{-1}(\varepsilon - M_m X_m). \quad (2.43)$$

This last situation corresponds to $n_X - m = q$, and X_{um} is estimated by setting $\varepsilon = 0$ and $X_m = Y$.

When $n_X - m > q$, there are more unmeasured variables than conservation constraints. The process is not fully observable (some states or even all states are non-estimable).

When $n_X - m < q$, there are more equations than unmeasured states. It is highly improbable that, due to the measurement uncertainties, a value of X_{um} that would simultaneously satisfy all the conservation constraints could exist. Since the number of equations in Ψ , $m + q$, is larger than the number of states to be estimated n_X , the observation system is redundant. The elimination of X_{um} from (2.42) leads to the linear redundancy equations:

$$RX_{mr} = T\varepsilon, \quad (2.44)$$

where X_{mr} is the vector of the redundant measured states. By replacing X_{mr} by the measurement Y_r , one obtains the values of the parity vector:

$$r = RY_r = Re_r + T\varepsilon. \quad (2.45)$$

This equation clearly shows that the redundancy equations residuals are not zero because of the uncertainties prevailing in the measurement and mass conservation constraints. To manage this problem, one possibility could be to remove redundant measurements. But this is the wrong approach, since experimental information is lost and, furthermore, the estimate values would depend upon the data that would have been removed. The right approach is to reconcile the information by the procedures discussed in this chapter.

Classification of the process variables: The condition $n_X - m \leq q$ is not sufficient to ensure process observability. It may happen that the process observability is only partial, *i.e.* that some states are non-observable. It may also happen that, though the system is not redundant, some variables are observable redundant, while, as a conse-

quence, others are non-observable. In fact any system of observation as depicted by Equation 2.32 or 2.35 can be decomposed into unobservable and observable parts. Moreover, the observable states can also be classified into minimal observability and redundant observability classes. Figure 2.6 shows the classification tree that can be drawn for the measured Z variables and the process states. The particular case where measured process variables are state variables ($Z = X_m$) is given within brackets.

The classification of the process variables and the decomposition of the equation system (2.32), or (2.35) in the linear case, into its redundancy part of Equation 2.37 are two related problems. The initial Equation 2.32 can be rewritten as

$$\begin{aligned} R(Z_r) &= 0, \\ Q(X_o) &= Q'(Z_r) + Q''(Z_{nr}), \\ X_{no}, \end{aligned} \quad (2.46)$$

where the second equation corresponds to the deductible part of the system, allowing calculation of the observable states X_o . Various methods are available for this decomposition of the initial system and the classification of the process variables. They have been developed mainly for linear systems, but also extended to multi-linear systems. The graph theory is used in [65] and [66]. The Gauss Jordan elimination technique is used in [67], projection matrices in [68] and [69], a mixture of graph theory and linear algebra in [70], and the QR factorization in [71]. Using adapted process variable combinations, these methods have been extended to bilinear systems and even to multi-linear systems [72].

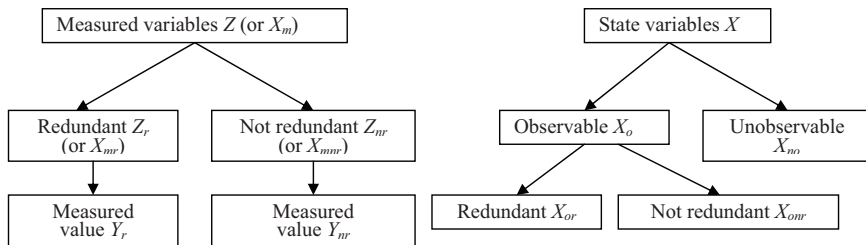


Figure 2.6 Scheme showing the status of the various process variables

2.6.2 General Principles for State Estimate Calculation

The data reconciliation problem to be solved is schematically represented in Figure 2.5, where the known information is given at the reconciliation procedure input and the process states to be estimated at the procedure output. When f and g are linear functions and inequality constraints are inactive, the solution to this problem is analytical. This occurs when the measured process variables are plant states or

linear combinations of them, and represent component flowrates. From the state estimates, one can subsequently correct the values of the measured variables and estimate the unmeasured process variables.

The optimization method that can be used to solve the unconstrained problem defined by (2.30) is the usual approach where the reconciled states are the \hat{X} values that cancel the derivatives of J with respect to X . The resolution methods are slightly different for the degenerate cases when either ε or e is zero.

When $\varepsilon = 0$, the optimization problem

$$\min_X J(X) \text{ subject to } f(X) = 0 \quad (2.47)$$

can be processed by one of two main approaches. The substitution method consists in exploiting the equality constraints to decrease the number of variables the criterion has to be minimized with respect to. The Lagrange method, on the contrary, increases the number of search variables to optimize the criterion by incorporating the equality constraints into it. New variables λ , called the Lagrange multipliers, are associated to these constraints.

- The substitution method consists of the following steps:
 1. Select a set of independent variables among X . Such a set is composed of the smallest number of variables X_{ind} that, if they were measured, would allow the remaining variables from the system $f(X) = 0$ (the dependent ones X_{dep}) to take unique values. In other words, the $n_X - q$ independent variables X_{ind} are selected in such a way that the q variables X_{dep} are at minimal observability.
 2. Express X_{dep} as a function of X_{ind} by solving

$$f(X_{ind}, X_{dep}) = 0 \quad (2.48)$$

with respect to X_{dep} :

$$X_{dep} = h(X_{ind}). \quad (2.49)$$

3. Replace X in the criterion J by its expression as a function of X_{ind} . The initial constrained minimization problem with respect to X is thus transformed into an unconstrained minimization problem with respect to X_{ind} :

$$\min_{X_{ind}} J(X_{ind}). \quad (2.50)$$

4. Find the \hat{X}_{ind} value and calculate \hat{X}_{dep} from (2.49).
- The Lagrange method consists in integrating the constraint $f(X) = 0$ into the J criterion to form a new function, the Lagrangian \mathcal{L} , which has to be optimum with respect to the variables X and λ , the latter being the Lagrange multipliers:

$$\mathcal{L} = J(X) + \sum_i \lambda_i f_i(X), \quad (2.51)$$

where f_i stands for the i th conservation constraint in the system $f(X)$, for $i = 1$ to q . The initial constrained problem has been replaced by a new optimization problem with $n_X + q$ unknown parameters.

These optimization problems can be solved by classical numerical techniques. When J is quadratic and the constraints are linear, an analytical solution can be obtained by resolving the system expressing that the derivatives of the reconciliation criterion, or the Lagrange function, have zero values. In other cases, either non-linear programming methods can be used to minimize the criterion, or numerical methods used for solving the non-linear equations expressing that the derivatives of the criterion are zero.

When $e = 0$, the optimization problem

$$\min_X J(X) = f(X)^T V_\varepsilon^{-1} f(X) \text{ subject to } g(X) = Y \quad (2.52)$$

can again be processed by one of the two above approaches: substitution and Lagrange methods. This reconciliation method is common practice in the bilinear case. It allows to estimate unmeasured flowrates through the use of species concentrations in the ore streams. It is called the node imbalance method.

2.7 The Linear Cases: Steady-state, Stationary and Node Imbalance Data Reconciliation Methods

2.7.1 The Steady-state Case

The unconstrained SSR problem can be formulated as the following particular case of (2.30):

$$\begin{aligned} \hat{X} &= \arg \min_X [(Y - CX)^T V^{-1} (Y - CX)], \\ Y &= CX + e; \quad e \sim N(0, V). \end{aligned} \quad (2.53)$$

In the substitution method, $n_X - q$ independent variables are selected using either heuristic or systematic methods, and the X vector is restructured as (X_{ind}, X_{dep}) , while the matrix M is partitioned into compatible blocks, such that the conservation constraints can be rewritten as

$$(M_{ind} \ M_{dep}) \begin{pmatrix} X_{ind} \\ X_{dep} \end{pmatrix} = 0. \quad (2.54)$$

or

$$M_{ind} X_{ind} + M_{dep} X_{dep} = 0. \quad (2.55)$$

The dependent variables X_{dep} can then be expressed as functions of X_{ind} :

$$X_{dep} = -(M_{dep})^{-1} M_{ind} X_{ind} \quad (2.56)$$

assuming that the independent variables have been selected to ensure M_{dep} invertibility. Therefore X can be expressed as

$$X = \begin{pmatrix} X_{ind} \\ X_{dep} \end{pmatrix} = \begin{pmatrix} X_{ind} \\ -M_{dep}^{-1}X_{ind} \end{pmatrix} = LX_{ind}, \quad (2.57)$$

with $L = \begin{pmatrix} I & -M_{dep}^{-1} \end{pmatrix}^T$.

The linear reconciliation criterion:

$$J = (CX - Y)^T V^{-1} (CX - Y) \quad (2.58)$$

can then be written as a quadratic expression of X_{ind} :

$$J(X_{ind}) = X_{ind}^T L^T C^T V^{-1} C L X_{ind} - 2Y^T V^{-1} C L X_{ind} + Y^T V^{-1} Y. \quad (2.59)$$

J is minimized with respect to X_{ind} by writing that the derivatives of J with respect to X_{ind} have zero values. One obtains a system of $n_X - q$ equations system with $n_X - q$ unknown states, which has the following solution:

$$\hat{X}_{ind} = (L^T C^T V^{-1} C L)^{-1} L^T C^T V^{-1} Y. \quad (2.60)$$

Knowing X_{ind} , the dependent variable estimates can be calculated from Equation 2.56:

$$\hat{X}_{dep} = -(M_{dep})^{-1} M_{ind} \hat{X}_{ind}. \quad (2.61)$$

The Lagrange method can also be applied to solve the linear SSR problem. The Lagrangian function is

$$\mathcal{L} = (CX - Y)^T V^{-1} (CX - Y) + \lambda^T M X. \quad (2.62)$$

The \mathcal{L} stationarity conditions are

$$\frac{d\mathcal{L}}{dX} = 2C^T V^{-1} CX - 2C^T V^{-1} Y + M^T \lambda = 0, \quad (2.63)$$

$$\frac{d\mathcal{L}}{d\lambda} = M X = 0. \quad (2.64)$$

The X solution of this $n_X + q$ equation system with $n_X + q$ unknown variables is

$$\hat{X} = \Pi^{-1} (I - M^T (M \Pi^{-1} M^T)^{-1} M \Pi^{-1}) C^T V^{-1} Y, \quad (2.65)$$

where

$$\Pi = C^T V^{-1} C. \quad (2.66)$$

There are other alternatives to the solution of the reconciliation problem. Two of them should be mentioned, in the particular case where the measured variables are states ($Z = X_m$):

- The simplest one, for software development, consists in selecting $C = I$, *i.e.*, assuming that all the states are measured. For states that are not actually measured, corresponding variances of measurement errors are simply given very large values. The solution is thus directly obtained from (2.65) as

$$\hat{X} = (I - VM^T(MVM^T)^{-1}M)Y. \quad (2.67)$$

This method is used in the BILMATTM algorithm [22], and has never numerically failed.

- The system of conservation constraints can be partitioned using the redundancy equations as follows (linear version of Equation 2.46):

$$\begin{cases} RX_{mr} = 0, \\ QX_o = Q'X_{mr} + Q''X_{mnr}. \end{cases} \quad (2.68)$$

The solution is then obtained in two steps. First, the reconciled values of the measured redundant states are estimated by

$$\hat{X}_{mr} = (I - V_{mr}R^T(RV_{mr}R^T)^{-1}R)Y_{mr}, \quad (2.69)$$

where V_r and Y_r are formatted to match with X_{mr} . Second, the following equation is solved for X_o :

$$Q\hat{X}_o = Q'\hat{X}_{mr} + Q''Y_{nr}. \quad (2.70)$$

Tuning of the measurement error matrix covariance. Before using the steady-state method, it is always necessary to test: (1) that the main contribution to the variability of the process variables is related to the measurement and sampling errors; and (2) that there is an underlying steady-state regime during the period of time corresponding to the measurement and sampling process. In this case, the measurement errors due to the process time variations, such as the integration error, should be incorporated into the measurement error. In other words data obtained during a transient regime of large magnitude in comparison with the measurement errors should not be used for SSR. Methods to find the measurement variance matrix have been discussed by [41],[73–78].

Four methods are proposed to estimate the V covariance matrix:

- Estimation from the properties of the equipment involved in the measurement process and the properties of the material flowing in the stream. For instance the theory of particulate material sampling [56, 79, 80] can be applied as seen in Section 2.4, in conjunction with a statistical analysis of the reliability of the measuring devices for the measured specific material property.
- Direct estimation of the variance of the measurement Y from a large set of data for the same plant operated under the same operating conditions. In this case, the estimate of V is obtained from the Y records by standard statistical estimation techniques, and is noted $Var(Y)$. This estimate of V does not strictly consist of measurement errors, since it also includes some process dynamic disturbances, such as the integration error. However, this is the right approach, since the recon-

ciled values are expected to give an estimate of the underlying plant steady-state (*i.e.*, average) behavior.

- Indirect estimation from the constraint residuals (the node imbalances) in the particular case where $C = I$ (the identity matrix). The constraint residuals r are calculated for each Y value of the data set and their variance estimated:

$$\begin{cases} r = MY = MX + Me, \\ \text{Var}(r) = MV M^T. \end{cases} \quad (2.71)$$

Techniques have been proposed [77] for extracting V from this last equation. The main advantage of this technique is that it takes account of the mass conservation constraints.

- Simultaneous estimation of V and of a rough model of the plant. For instance, a mineral separation plant can be simply modeled by mineral separation coefficients at each separation node of the flowsheet. The plant model is then expressed as

$$X = B(s)X_f \quad (2.72)$$

where X_f is the state vector of the feed streams and s the separation coefficient vector. The variance of Y estimated from a measurement data set is then

$$\text{Var}(Y) = B(s)\text{Var}(X_f)B(s)^T + V. \quad (2.73)$$

From Equation 2.73, variance V can be extracted simultaneously to s and $\text{Var}(X_f)$ by a least-squares procedure [78].

2.7.2 The Stationary Case

The unconstrained stationary reconciliation problem is formulated as the following particular case of (2.30):

$$\begin{cases} \hat{X} = \arg \min_X [(Y - CX)^T V^{-1} (Y - CX) + \varepsilon^T V_\varepsilon^{-1} \varepsilon], \\ Y = CX + e; e \sim N(0, V), \\ \varepsilon = MX; \varepsilon \sim N(0, V_\varepsilon). \end{cases} \quad (2.74)$$

The solution is obtained by canceling the derivatives of J :

$$\frac{dJ}{dX} = 2C^T V^{-1} CX - 2C^T V^{-1} Y + 2M^T V_\varepsilon^{-1} MX = 0, \quad (2.75)$$

which leads to

$$\hat{X} = (C^T V^{-1} C + M^T V_\varepsilon^{-1} M)^{-1} C^T V^{-1} Y \quad (2.76)$$

or alternatively, using a matrix inversion lemma,

$$\hat{X} = \Pi^{-1} [I - M^T (V_\varepsilon + M \Pi^{-1} M^T)^{-1} M \Pi^{-1}] C^T V^{-1} Y, \quad (2.77)$$

where Π is defined in Equation 2.66. Equation 2.77 gives Equation 2.65 when V_ε is set to 0.

Tuning of the accumulation rate covariance matrix. Three methods are proposed to estimate the V_ε covariance matrix:

- Measurement of the state variables by a suitable combination of instrumentation for a sufficiently long period of time. This may require transient installation of instruments on the streams. Since it may happen that not enough sensors are available to simultaneously measure all the state variables, the different parts of the circuit may be sequentially studied, assuming that the statistical properties of the operation do not change much from a test on a plant sub-network to a test on another sub-network. For such a subset of X , the measurement equation is

$$Y = X + e. \quad (2.78)$$

Multiplying Equation 2.78 by the incidence matrix M of the sub-network corresponding to the fully instrumented part of the plant, one obtains

$$MY = MX + Me = \varepsilon + Me. \quad (2.79)$$

Assuming that there is no correlation between the state variables and their measurement errors, the variance-covariance matrix is given by:

$$MVar(Y)M^T = V_\varepsilon + MVM^T, \quad (2.80)$$

where $Var(Y)$ is the measurement variance-covariance matrix of Y , the elements of which can be evaluated by the usual statistical estimation methods. V_ε can then be calculated by

$$V_\varepsilon = MVar(Y)M^T - MVM^T. \quad (2.81)$$

The total V_ε matrix can therefore be constructed by consistently assembling the various terms of the partial variance matrices. Obviously some covariance terms will not be evaluated, but since they necessarily correspond to distant streams in the plant, they can be neglected.

- Alternatively, one may evaluate the V_ε matrix by an approximation of the plant dynamics, using for instance first order transfer functions with rough evaluation of their time constants. The process dynamics are subsequently gathered into a plant state-space model (see Section 2.10 for more detailed explanations). This model allows an estimation of V_x , the state variable variance characterizing the underlying unknown random dynamic variations of the process. V_ε can finally be estimated by

$$V_\varepsilon = MV_xM^T. \quad (2.82)$$

- A third option is to use V_ε as a tuning factor that can be adjusted heuristically following an evaluation of the data reconciliation performance in comparison with a desired behavior. In fact, this is the subjective method that is mostly adopted, for example, when tuning model uncertainty relative variances and measurement uncertainties in Kalman filtering techniques.

Example. Figure 2.7 shows a standard flotation circuit arrangement.

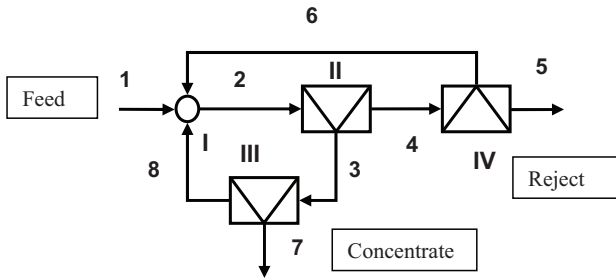


Figure 2.7 A flotation mineral separation circuit

The state vector X consists of the valuable mineral flowrates in the eight streams. There are four nodes, two recycling streams and two plant outputs, the concentrate, which contains the valuable mineral to be sold for metal extraction, and the reject. Table 2.2 gives the measured values Y of the flowrates of the mineral to be concentrated, as well as their reliability σ_e (standard deviations of e). Table 2.2 shows also the values of the reconciled values \hat{X} . The reconciled values do not exactly satisfy the steady-state mass conservation equations since the plant is assumed to be operated in a stationary dynamic regime where V_e is different from zero.

Table 2.2 Data for the circuit of Figure 2.7 (fifth column will be explained in Section 2.9)

Stream	$Y(t/h)$	σ_e	\hat{X}	$\sigma_{\hat{x}}$
1	15,8	0,6	14,86	0,44
2	22,2	0,8	23,08	0,45
3	-	∞	14,19	0,47
4	8,3	0,3	8,18	0,27
5	1,4	0,2	1,47	0,19
6	6,3	0,3	6,23	0,23
7	11,3	0,5	11,62	0,46
8	2,5	0,2	2,45	0,2

2.7.3 The Node Imbalance and Two-step Methods for Bilinear Systems

In bilinear constraint cases, when either flowrates or species concentrations are assumed to be exactly measured, the reconciliation procedure degenerates into a LQ

(linear-quadratic) optimization problem. The node imbalance for the component i is defined as the mass flowrate residuals at the network nodes that are obtained when using the measured values of the component i concentrations:

$$\varepsilon_i = M_i(f_0 \bullet c_i) = M_i C_i^d f_0, \quad (2.83)$$

where C_i^d is the diagonal matrix of c_i . In the node imbalance method, it is assumed that component concentrations are known, and that flowrates only have to be estimated. The optimal flowrate estimates are those that minimize

$$J(f_0) = \sum_{i=1}^n \varepsilon_i^T V_{\varepsilon_i}^{-1} \varepsilon_i + (Y_f - C_f f_0)^T V_f^{-1} (Y_f - C_f f_0), \quad (2.84)$$

where the first term of J corresponds to the node imbalances and the second term to the flowrate measured values, if any:

$$Y_f = C_f f_0 + e_f; \quad e_f \sim N(0, V_f). \quad (2.85)$$

The node imbalance estimates of the flowrates are therefore

$$f = (\Sigma_i \gamma_i + C_f^T V_f^{-1} C_f)^{-1} C_f^T V_f^{-1} Y_f, \quad (2.86)$$

where

$$\gamma_i = C_i^{dT} M_i^T V_{\varepsilon_i}^{-1} M_i C_i^d. \quad (2.87)$$

When the flowrates have been estimated by the node imbalance method, species concentrations can be estimated in a second step. When the flowrates are assumed to be known, the species concentration constraints become linear, as shown in this expression:

$$M_i(f_0 \bullet c_i) = M_i F_0^d c_i = 0, \quad (2.88)$$

where F_0^d is the diagonal matrix of the known f_0 values. The quadratic criterion to be minimized, subject to (2.88), is

$$J(\dots, c_i, \dots) = \sum_{i=1}^n (Y_i - C_i c_i)^T V_i^{-1} (Y_i - C_i c_i), \quad (2.89)$$

where C_i is the observation matrix of the component i concentration, and V_i its corresponding measurement error variance matrix. The solution is given by the same equations as (2.65) and (2.66), providing that there are no additional constraints that couple the mass balance equations of the various components:

$$\hat{c}_i = \alpha_i [I - F_0^d M_i^T (M_i F_0^d \alpha_i F_0^d M_i^T)^{-1} M_i F_0^T \alpha_i] C_i^T V_i^{-1} Y_i, \quad (2.90)$$

where

$$\alpha_i = (C_i^T V_i^{-1} C_i)^{-1}. \quad (2.91)$$

When the node imbalance estimation of the flowrates is followed by the reconciliation of the species concentrations for these flowrate estimates, the method is a sub-optimal two-step LQ method [34]. The next section shows that a similar hierarchical technique can be used to find the true optimal solution to the bilinear reconciliation problem.

2.8 The Non-linear Cases

As discussed above, when the constraints and the measurement equations are linear, the solution of the reconciliation problem can be developed analytically. However, in non-linear cases, it is normally impossible to derive an explicit expression for the reconciled states. Several methods (for instance [26], [81–86]) are possible depending upon the approach selected to handle the constraints (substitution or Lagrange multipliers techniques), and the optimization technique that is used to minimize the criterion. The following optimization techniques are possible options:

- Any Non-linear programming (NLP) method, involving a search algorithm to iteratively approach the optimal values. The substitution method is well adapted to these procedures since it decreases the number of search variables to be manipulated by the NLP algorithm.
- A hierarchical minimization method where the criterion and the search variables are split into blocks. These approaches allow hybrid minimization methods, in which some parts are optimized by analytical methods, others by NLP algorithms.
- Iterative numerical resolution of the system of equations expressing that the Lagrange criterion derivatives have zero values.

It would be too long to detail all the possible options for solving non-linear reconciliation problems, but examples may help to illustrate some possible approaches.

2.8.1 *An Example of Substitution Methods: Mass and Heat Balance of a Thermal Exchanger*

Let us consider the mixer-exchanger of Figure 2.8 [87], which heats an electrolyte solution containing dissolved copper and nickel subsequently fed to an electro-refinery plant. Part of the solution is directly recycled to the mixer-exchanger, and another part is cleaned from impurities before being recycled. Data collected and averaged for a time period of 3 h are presented in Table 2.3, together with an estimation of their standard deviations. The process variables are the solution flowrates F , densities ρ , temperatures T , copper and nickel concentrations c and n for the streams 1, 2 and 3, and heating power P . The system parameters are the specific heats of the solutions C , respectively 4400, 4210, 4350 J/(kg.K) for streams 1, 2

and 3. They are assumed to be constant and perfectly known. The data are assumed to reasonably represent a steady-state operating regime of the plant. The thermal losses are neglected.

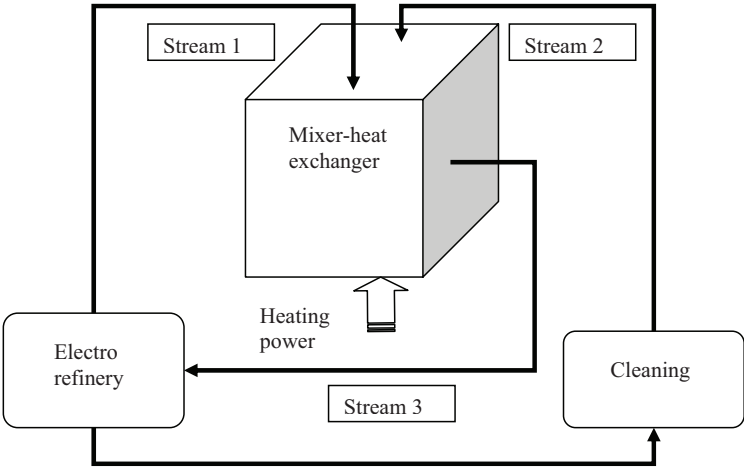


Figure 2.8 Mixer-exchanger for conditioning an electrolyte

Table 2.3 Data for the plant of Figure 2.8

Variable	T (°C)			F (m ³ /min)			ρ (kg/l)			c(g Cu/l)			n(g Ni/l)		
	Y	σ	\hat{X}	Y	σ	\hat{X}	Y	σ	\hat{X}	Y	σ	\hat{X}	Y	σ	\hat{X}
1	46	2	45.3	20	4	20.4	1.31	0.12	1.25	40	3	41.1	10	2.0	10.1
2	27	2	26.7			9.6	1.15	0.06	1.15	10	2	10.2	1	0.3	1.0
3	42	3	42.1			27.2	1.28	0.12	1.34	37	4	34.4	8	2.0	7.9
Power data (kJ/s) P : $Y = 6000$; $\sigma = 500$; $\hat{X} = 5978$															

The mass balance constraints around the mixer-exchanger are written as

$$\begin{cases} F_1\rho_1 + F_2\rho_2 - F_3\rho_3 = 0 \\ F_1c_1 + F_2c_2 - F_3c_3 = 0 \\ F_1n_1 + F_2n_2 - F_3n_3 = 0 \end{cases} \tag{2.92}$$

and the heat balance as

$$F_1\rho_1C_1T_1 + F_2\rho_2C_2T_2 - F_3\rho_3C_3T_3 + P = 0 \tag{2.93}$$

The least-squares criterion is

$$J(X) = \sum_{i=1}^3 \left[\frac{(T_i - T_i^m)^2}{\sigma_{T_i}^2} + \frac{(c_i - c_i^m)^2}{\sigma_{c_i}^2} + \frac{(n_i - n_i^m)^2}{\sigma_{n_i}^2} + \frac{(\rho_i - \rho_i^m)^2}{\sigma_{\rho_i}^2} \right] + \frac{(F_1 - F_1^m)^2}{\sigma_{F_1}^2} + \frac{(P - P^m)^2}{\sigma_P^2}, \quad (2.94)$$

where the superscript m stands for the measured value of the process variable (Y) and the subscript i for the stream number. The minimization of $J(X)$ subject to the bilinear and trilinear constraints (2.92) and (2.93) can be performed with a PNL algorithm applied to the substitution method. The independent variables are selected as

$$X_{ind} = (F_1 \ F_2 \ F_3 \ T_1 \ T_2 \ T_3 \ \rho_1 \ \rho_2 \ c_1 \ c_2 \ n_1 \ n_2)^T. \quad (2.95)$$

The dependent variables are then

$$X_{dep} = (\rho_3 \ c_3 \ n_3 \ P)^T. \quad (2.96)$$

The minimization algorithm is depicted in Figure 2.9. It requires an initialization of the independent variables, which can be estimated from the conservation constraints and the measured values, using, for instance, the node imbalance method.

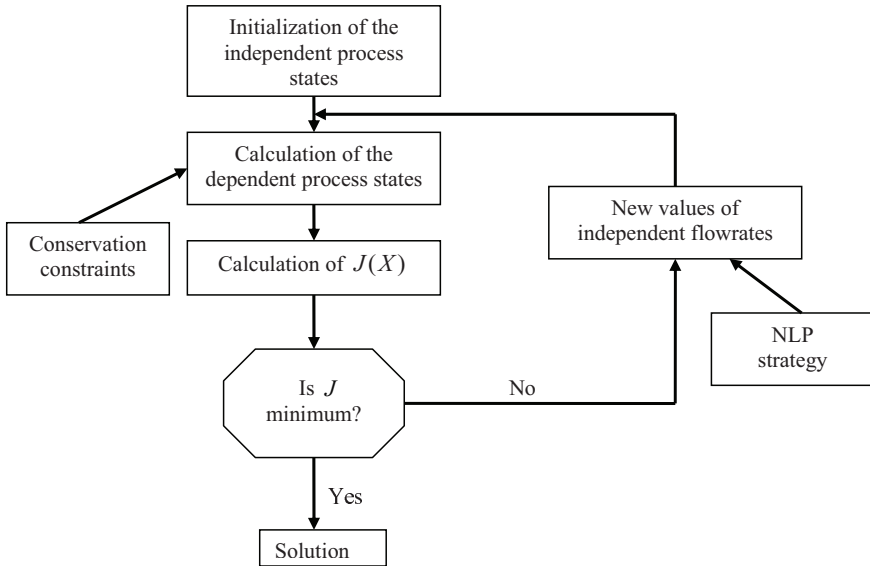


Figure 2.9 Algorithm for the calculation of the reconciled values by the substitution/PNL method

The results of the reconciliation procedure appear in Table 2.3. The simultaneous negative and positive corrections of, respectively, power and stream 3 temperature indicate that thermal losses should be taken into account. Repetition of the measurement campaign as well as application of FDI techniques (see Section 2.12)

could confirm this assumption. In that case a model for estimating the losses could be used to generate a virtual measure of the heat loss that can be added to the reconciliation problem.

2.8.2 An Example of Hierarchical Methods: BILMAT™ Algorithm

The BILMAT™ algorithm has been developed mainly for bilinear problems [22] but it can be extended to trilinear systems [53, 59]. The basic principle is that, when the main phase flowrates are known (upper level of the hierarchy), the species mass conservation constraints become linear functions of the species mass fractions (see Section 2.7.3). Therefore the idea is to define a lower level reconciliation problem where the criterion contains the mass fraction part of the data, and the constraints are written only for the species conservation. This problem, being LQ, has a direct analytical solution (see Section 2.7.3). At the upper level, only the flowrate variables are manipulated to minimize the overall criterion. Formally the problem is thus split into two minimization problems defined by

$$\min_{X_f, X_m} J(X_f, X_m) = \min_{X_f} \left(\min_{X_m} J_m(X_m | X_f) + J_f \right), \quad (2.97)$$

where X_f and X_m are, respectively, the flowrate state variables and the species mass fractions variables, and J_m and J_f the J parts containing the measured mass fractions and flowrates, respectively. Since the upper minimization level is non-linear, one can apply the substitution/PNL method to decrease the number of search variables. This method is depicted in Figure 2.10.

To illustrate the method, let us again consider the introductory example of Section 2.1.1. Flowrates F_1 and F_2 can be selected as independent flowrates of the upper level, since F_3 can be deduced from

$$F_3 = F_1 - F_2. \quad (2.98)$$

Giving values to these three variables leads to a linear structure for the copper and zinc mass conservation equations:

$$\begin{aligned} F_1 c_1 - F_2 c_2 - F_3 c_3 &= 0; \\ F_1 z_1 - F_2 z_2 - F_3 z_3 &= 0. \end{aligned} \quad (2.99)$$

The calculation of the reconciled metal mass fractions is then a LQ problem given the values of the flowrates. The solution to this problem is obtained by forming the Lagrange function:

$$\mathcal{L} = J_c + 2\lambda_c(F_1 c_1 - F_2 c_2 - F_3 c_3) + 2\lambda_z(F_1 z_1 - F_2 z_2 - F_3 z_3), \quad (2.100)$$

where J_c is the part of the criterion which contains the metal assays:

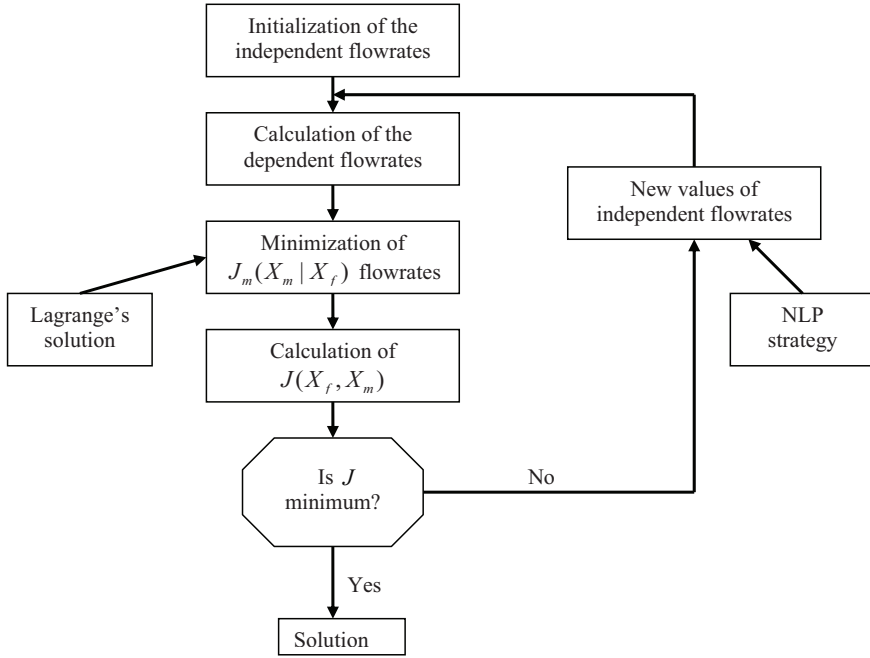


Figure 2.10 BILMAT™ reconciliation algorithm

$$J_c = \sum_{i=1}^3 \frac{(c_i - c_i^m)^2}{\sigma_{c_i}^2} + \sum_{i=1}^3 \frac{(z_i - z_i^m)^2}{\sigma_{z_i}^2}. \quad (2.101)$$

Then the Lagrange function is derived with respect to the six metal mass fractions and the two Lagrange multipliers, and the derivative equations set to 0. The solution of this system of eight unknown variables into eight equations leads to

$$\begin{aligned} \hat{c}_1 &= c_1^m - \lambda_c F_1 \sigma_{c_1}^2 \\ \hat{c}_2 &= c_2^m - \lambda_c F_2 \sigma_{c_2}^2 \\ \hat{c}_3 &= c_3^m - \lambda_c F_3 \sigma_{c_3}^2 \end{aligned} \quad (2.102)$$

with

$$\lambda_c = \frac{F_1 c_1^m - F_2 c_2^m - F_3 c_3^m}{F_1^2 \sigma_{c_1}^2 + F_2^2 \sigma_{c_2}^2 + F_3^2 \sigma_{c_3}^2} \quad (2.103)$$

for copper, and similar expressions for zinc. The reconciled values obtained from Equations (2.102) are optimum only for the selected particular values of the flowrates. To minimize the overall criterion

$$J = J_c + \sum_{j=1}^3 \frac{F_j - F_j^m}{\sigma_{F_j}^2}, \quad (2.104)$$

a NLP is required to find the optimal values of F_1 and F_2 , as shown in Figure 2.10. The optimal values are given in Table 2.1.

This hierarchical technique has been extended to trilinear and quadrilinear problems, *i.e.*, problems involving slurry, ore and water, particle size distributions, and chemical assays within particle size classes. Sub-optimal solutions have been applied to investigate the path followed by precious metals in a base metal sulphide plant [59], in a uranium ore grinding circuit [59], and in gold ore comminution and leaching circuits [28].

2.9 Performance of Data Reconciliation Methods

When the reconciliation calculation is achieved, it is worthwhile to estimate the reliability of the results in comparison with those of the raw measurements. This is performed by a sensitivity analysis of the propagation of the measurement errors and rate of accumulation uncertainties through the reconciliation process [34] and [44, 88]. In other words, the variance–covariance matrix of the reconciled states has to be calculated for an assessment of the reconciliation procedure benefits.

There are two approaches to calculate the variance matrix of the reconciled states, $V_{\hat{X}}$, *i.e.*, the variance of the estimation error. One can explicitly formulate the reconciled states as functions of Y and then calculate the variance $V_{\hat{X}}$ as a function of V and V_e . When the reconciliation method leads to a linear estimator, *i.e.*, when the reconciled states are linear functions of Y , then the usual linear algebra of variance calculation can be applied. Otherwise, one can linearize the expression around the process steady-state values and apply linear variance algebra. The second method consists in randomly generating synthetic values of Y according to a normal distribution $N(\hat{Y}, V)$ and repeating the reconciliation procedure. If this is done a sufficiently large number of times, the statistical properties of \hat{X} can be subsequently estimated. This method is known as a Monte-Carlo sensitivity analysis.

When the X estimator is linear, one can calculate the variance of \hat{X} using the relationship between \hat{X} and Y . For instance, in the linear steady-state case, Equation 2.65 gives

$$\hat{X} = \Gamma Y, \quad (2.105)$$

where Γ is given by

$$\Gamma = \Pi^{-1} [I - M(\Pi \Pi^{-1} M^T)^{-1} M \Pi^{-1}] C^T V^{-1}. \quad (2.106)$$

Therefore \hat{X} variance is

$$V_{\hat{X}} = \Gamma V \Gamma^T. \quad (2.107)$$

Since \hat{X} , in the assumed Gaussian context, is a maximum likelihood estimate, it is obvious that this estimate is such that necessarily the diagonal terms of $V_{\hat{X}}$ are lower than the diagonal terms of the variance of reconciled values obtained by any other estimator. It has been shown [44] that the variance of the reconciled measured

variables is

$$V_{\hat{Y}} = V - V_{in}, \quad (2.108)$$

where V_{in} is the variance of the innovations in , which are residuals of the observer, *i.e.*, the differences between the measured and reconciled values:

$$in = Y - \hat{Y} = Y - C\hat{X}. \quad (2.109)$$

Since V_{in} has, by definition, positive diagonal terms, one obtains

$$V_{\hat{Y}}(i, i) \leq V(i, i) \text{ for } i = 1 \text{ to } m, \quad (2.110)$$

an inequality showing the variance reduction due to data reconciliation. A global variance reduction factor ρ_v can be defined and calculated for the case where a subset of states is measured ($Z = X_m$):

$$\rho_v = \frac{1}{m} \sum_{i=1}^m \frac{V(i, i) - V_{\hat{Y}}(i, i)}{V(i, i)} = \frac{1}{m} [q + m - n_X]. \quad (2.111)$$

This equation shows that the reduction factor is directly related to the degree of redundancy ρ_r of the information on the system (see Equation 2.34). The expression (2.111) is valid also for bilinear SSR problems, providing that the mass conservation constraints are linearized around the reconciled states [44].

In addition to this difference between the variances before and after reconciliation, there is a strong difference that appears in the covariance terms. It is here assumed that the measurement errors are not correlated. However, the estimation errors are now correlated because of the correlation induced by the conservation equations. Thus, matrices $V_{\hat{X}}$ and $V_{\hat{Y}}$ exhibit covariance terms that correspond to the state consistency induced by the reconciliation procedure. This last property has a strong impact on the reliability of the process performance indicators or process model parameters subsequently calculated using reconciled states instead of raw measurements [89, 90]. When the stationary reconciliation method is used rather than the steady-state method, the covariance terms are usually smaller since the conservation equations are given more flexibility by relaxing these constraints through the tuning of V_ϵ .

Example 1. The impact of the variance reduction and the role of the covariance terms have been illustrated for the calculation of separation plant recoveries in [89]. Figure 2.11 shows a separation node whose separation efficiency of component i can be calculated by one of the two possible formulae:

$$R_i = \frac{F_2 x_{2i}}{F_1 x_{1i}}, \quad (2.112)$$

$$R_i = \frac{x_{3i} - x_{1i}}{x_{3i} - x_{2i}} \frac{x_{2i}}{x_{1i}}. \quad (2.113)$$

This last formula is known as the two-product formula. When using raw data, one obtains quite different results depending on the expression used to calculate the

efficiency (for complex plants and data sets there is a very large number of paths to calculate separation efficiencies or recoveries, Equations 2.112 and 2.113 being just two simple examples). A major advantage of using reconciled data is that there is a unique efficiency value whatever the formula used for calculation. Moreover, its estimate is more reliable.

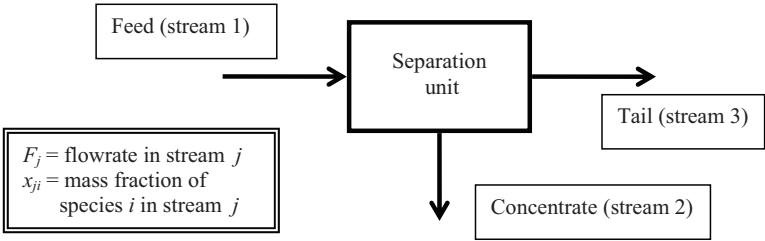


Figure 2.11 A separation unit

Example 2. Let us consider again the flotation circuit of Figure 2.7. Table 2.4 compares the standard deviations of the experimental and reconciled values of the mineral flowrates. The results show that reconciled data is more accurate than raw data. In addition to this positive effect and to the estimation of unmeasured variables, a drastic improvement of the process performance evaluation is usually brought on by the reconciled data. In this example, the relative standard deviation of the metal recovery (ratio of metal in the concentrate to metal in the feed) goes from 15% down to 6%, an improvement that would allow much more efficient decision for plant optimization.

Table 2.4 Reliability of measured and reconciled values (error standard deviations) for the plant of Figure 2.7

Stream	s.d of measured value	s.d. of reconciled value
1	0.6	0.44
2	0.8	0.45
3	∞	0.47
4	0.3	0.27
5	0.2	0.19
6	0.3	0.23
7	0.5	0.46
8	0.2	0.2

Example 3. Table 2.5 shows the results obtained for part of a copper–lead–zinc flotation plant as depicted in Figure 2.12 [90]. Clearly, when using raw data, one finds different values of the recoveries depending upon the formula used. Furthermore some calculations are very unreliable (standard deviations of 31 and 38%). The role

of the $V_{\hat{x}}$ variance and covariance terms is considerable in the drastic increase of the precision when using reconciled data.

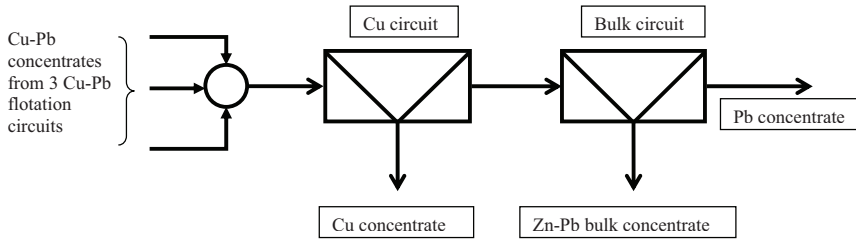


Figure 2.12 Copper–lead–zinc flotation plant

Table 2.5 Recoveries R and their reliabilities (standard deviations σ_R) as calculated with reconciled and raw data

Method	Steady-state reconciliation		Raw data and (2.112)		Raw data and (2.113)	
Recovery of	$R(\%)$	$\sigma_R(\%)$	$R(\%)$	$\sigma_R(\%)$	$R(\%)$	$\sigma_R(\%)$
Copper in copper concentrate of copper circuit	78.0	4.0	59.4	31	79.8	7.2
Lead in lead concentrate of bulk circuit	95.3	3.0	89.3	13	79.0	38
Zinc in bulk concentrate of bulk circuit	56.0	7.0	29.6	13	77.8	10

Example 4. Another use of reconciled data is the estimation of model parameters [91] such as grinding, leaching, and flotation rate constants [58], and hydrocyclone selectivity. For example, the confidence intervals of the estimated rates of breakage in an industrial ball mill of a closed circuit are presented in Figure 2.13 as functions of particle size [90]. Using reconciled values instead of raw data drastically improves the model reliability, particularly in the fine size range where the sensitivity of the rate of breakage is extremely high, because of the typical structure of the breakage phenomena, which involves chain reactions.

2.10 An Overview of Dynamic Reconciliation Methods

Steady-state data reconciliation is frequently used to enhance process information quality. However steady-state operating conditions do not exist in practice and corresponding methods cannot fulfill the need for reliable dynamic state estimation,

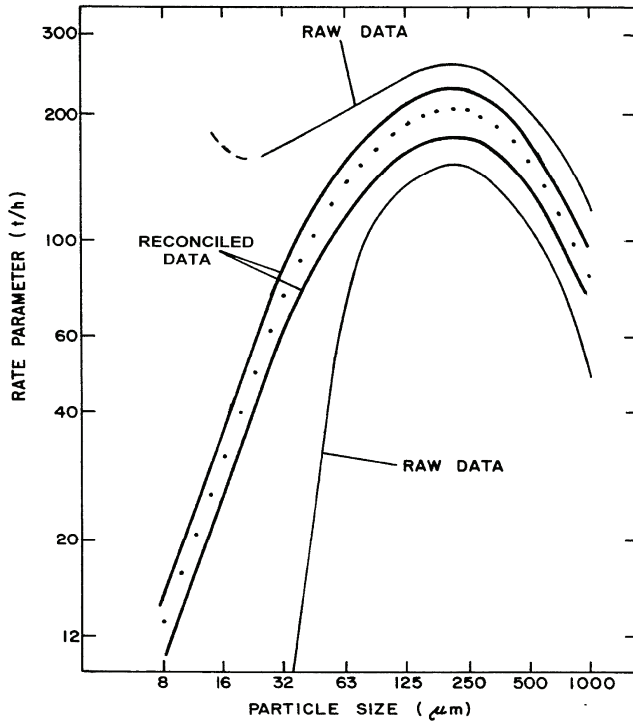


Figure 2.13 Standard deviation intervals for grinding rate parameters calculated from either raw or reconciled data

particularly when the state variable amplitude variations due to process dynamics are larger than the inherent measurement errors.

Various options are available to tackle the dynamic data reconciliation problem [40]. On-line application of SSR is the first option [54, 92], but it usually generates estimation errors larger than the measurement errors [76]. Averaging techniques coupled to steady-state data reconciliation can be used to compensate dynamic variations. However, this usually induces phase lags that deteriorate the process control performances [93], and tuning problems of the time window [65, 94]. Dynamic reconciliation is obviously required for batch or semi-batch processes [31, 95].

Mass conservation constraints incorporating material accumulation terms can be used [36, 37], but usually metallurgical inventories are not measurable or extremely inaccurate. Furthermore, the observability of these inventory variables, when unmeasured, is quite problematic. Treatment of accumulation rates as stochastic elements, as in the stationary method presented in Section 2.7.2, is a powerful alternative to the explicit use of inventories. However, the tuning of the variance of the uncertainties related to the neglected dynamics is a difficult task [96] for on-line application of stationary reconciliation methods [97].

Full model observer (FMO) based on phenomenological or empirical [121] process models can also be constructed, but the parameters are frequently so inaccurate and time varying that the observer performances are poor, therefore justifying the use of sub-models, more reliable than full models [98]. A number of intermediate methods between the SSR procedure and the full model-based observer have been experimented with and deserve to be clarified, since they are, in practice, usually more efficient than SSR or FMO approaches.

The objective of this section is to make a short survey of the possible dynamic data reconciliation methods, classified according to the models and reconciliation criteria used.

Two main classes of data reconciliation models will be considered: full (or causal) models, *i.e.*, models that can simulate process states and outputs from input variables and initial states, and sub-models that express mass or energy conservation constraints to be obeyed by the process variables. In both types of models, uncertainties and disturbances are considered since the observers to be designed must have the ability to process the uncertainties about the process behavior. The full model can be phenomenological or empirical, or a mixture of both approaches. The minimal sub-model necessarily contains phenomenological mass conservation constraints, but it can be augmented by other modeling considerations that may be phenomenological or empirical.

2.10.1 Phenomenological Causal Models

Although phenomenological models are obviously specific to a given metallurgical process, they nevertheless contain ingredients that are common to most chemical reactors (mixing and transport properties, reaction kinetics and equilibrium, material or energy balance equations). The flowing material is divided into $n + 1$ components i ($i = 0$ to $n + 1$), and the process variables are the component flowrate vectors f_i , in the p plant streams, stacked into the vector f , and the n_n unit process inventory vectors m_i . The mass conservation constraints are then (see Equation 2.6 where the structural uncertainty terms have been omitted):

$$\frac{dm_i}{dt} = M_i f_i + P_i \quad i \in \{0 \dots n\}, \quad (2.114)$$

where M_i is the incidence matrix for species i , and P_i the vector of production rates. The flowrates and production rates are usually related to the equipment inventories through kinetic parameters (such as breakage, flotation, leaching, reducing rate constants), and through the mixing and material transport properties:

$$\psi f = \varphi(m), \quad (2.115)$$

where ψ is a matrix that extracts the node output streams, and φ expresses the relationship between the flowrate and the equipment load. Finally, the disturbances or

model uncertainties must be modeled. The plant feed stream variations are modeled by the following linear stochastic equations:

$$z(t+1) = Az(t) + \xi(t), \quad (2.116)$$

$$\Omega f(t) = Bz(t) + \Omega \bar{f}, \quad (2.117)$$

where A is a state matrix, $\xi(t)$ a white noise of variance V_f , Ω the matrix that extracts the component feed rates from the plant stream flowrates stacked into f , \bar{f} the mean value of f , and B the matrix defining the feed rates from the state variable z . Parameter and model disturbances in Equations 2.114 and 2.115 are represented by additive uncertainties $\delta(t)$ that can be modeled by equations similar to Equations 2.116 and 2.117 driven by a white noise $\xi_\delta(t)$ of variance V_δ .

The model described by Equations 2.114 to 2.117 is a causal model since it can be used to simulate the process from the known inputs $\Omega \bar{f}$, $\xi_\delta(t)$, and $\xi(t)$. However, it contains a very large number of parameters (rate constants and feed coefficients, parameter and model disturbances), and is non-linear. An alternative to this approach is to construct empirical causal models.

2.10.2 Empirical Causal Model

To illustrate an example of an empirical model, let us consider any mineral processing plant flowsheet, linearized around nominal stream flowrates \bar{f} . It can be modeled by a network of connected basic units including (see Figures 2.14 and 2.15):

- stream junctions, *i.e.*, zero dynamics elements that combine flow streams;
- stream separators, *i.e.*, zero dynamics elements characterized by separation coefficients s_{ki} representing the split of component i between the two products of node k ;
- unit gain transfer functions (usually pure delays or low order transfer functions) representing the dynamic relationships that exist between the node feed and its output streams;
- feed disturbance generators driven by white noise signals ξ_i , as considered in Equation 2.116;
- internal disturbance generators, driven by white noise $\xi_s(t)$ used to model the separation efficiency disturbances. They are designed to preserve material conservation by adding to one stream the same amount of material that is removed from the other one.

The various connected elements can be gathered into a generic linear state space model:

$$z(t+1) = A'(t) + \xi'(t), \quad (2.118)$$

$$f(t) = B'z(t) + \bar{f}. \quad (2.119)$$

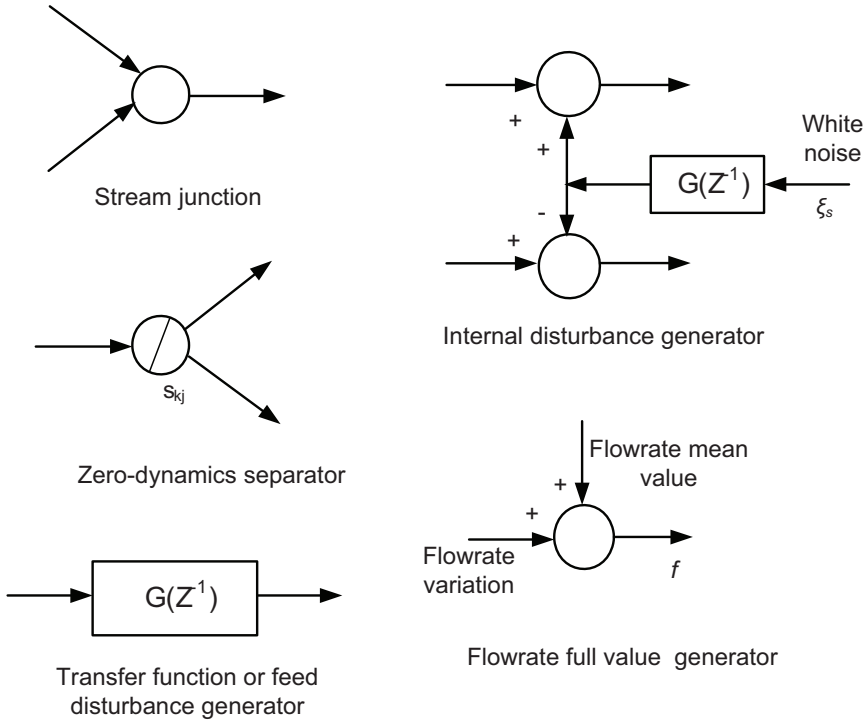


Figure 2.14 Basic elements of an empirical model of a mineral separation plant

This empirical full model implicitly contains the mass conservation Equations 2.8, in the absence of component transformation as assumed in the operating conditions defined by the previous selected basic units. Mass conservation for component i is written

$$\frac{dm_i}{dt} = M_i f_i = M_i (f_i - \bar{f}_i) = M_i B'_i z(t) \quad \text{for } i = 0 \text{ to } n + 1 \quad (2.120)$$

since the constraint $M_i \bar{f}_i = 0$ is valid for the nominal state values. A representation of the model for a single flotation cell is given in Figure 2.15. The same structure can be used for all the stream components and for all the nodes of a complex plant.

The model requires, in addition to the steady-state values \bar{f} , the knowledge of the separation coefficients s_{kj} , the transfer function coefficients, and the variance of the driving white noise vector ξ' including ξ_i and ξ_s . Both the empirical and the phenomenological models contain numerous parameters to be calibrated. This is why it is interesting to look at sub-models for dynamic data reconciliation.

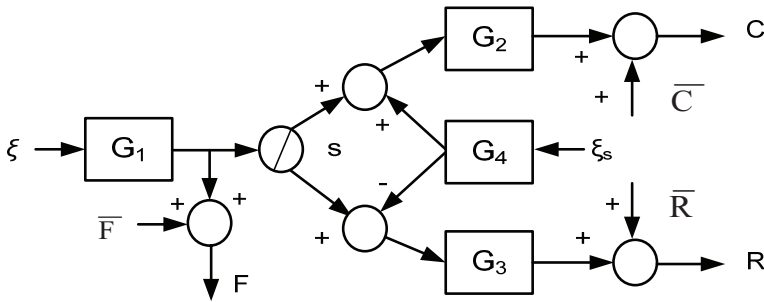


Figure 2.15 An empirical model of a flotation single unit (F, R and C stand for feed, concentrate and reject streams, and G for transfer functions, s is species separation coefficient, and ξ_s are white noise)

2.10.3 Sub-models

The minimal set of equations that can be used for designing model-based data filtering methods is the mass conservation constraints of Equations 2.114 written for components that are not transformed:

$$\frac{dm_i}{dt} = M_i f_i \quad i \in \{0 \dots n+1\} \quad (2.121)$$

or

$$\frac{dm}{dt} = Mf \quad (2.122)$$

using the stacked vectors f and m , and the block-diagonal matrix of the M_i s. This equation does not contain any parameters to be calibrated or uncertainties to be estimated. Obviously these would be ideal conditions for data reconciliation. Unfortunately the number of equations $q (= n_n(n+1))$ is low compared with the number $n_X (= (p+n_n)(n+1))$ of variables to be estimated. Hence – even if all the process variables are measured – this creates a low redundancy level, which in turn leads to limited improvements of the process variable reliability. In addition, the inventories m_j are usually not available for measurement.

To cope with the problem of inventory unavailability, it is possible to consider the stationary model described in Section 2.7.2 for instantaneous data reconciliation, while adding time correlation through the statistical behavior of the accumulation rates (node imbalances):

$$Mf = \varepsilon \quad \varepsilon \sim N(0, V_\varepsilon(\tau)), \quad (2.123)$$

where τ is the time lag used for calculating $V(\tau)$ which is the autocovariance of ε . This stationary dynamic model degenerates into the stationary model of Section 2.7.1 when $V_\varepsilon(\tau) = 0$ for $\tau > 0$, or into the steady-state mass balance when $V_\varepsilon = 0$.

To add time links between successive values of stream variables included in Equations 2.122, it has been proposed to add empirical evolution models for each flowrate [36]. Random walk, for instance, can be used:

$$f(t+1) = f(t) + \omega(t) \quad \omega \sim N(0, V_\omega), \quad (2.124)$$

where V_ω is considered to be a diagonal matrix. This model is rather a rough and arbitrary approximation since it does not take account of the physical links that must exist between the input and output flows of a process. Stationary autoregressive models should be preferred for describing the time evolution of the variables.

Instead of adding empirical information to the mass conservation constraint of Equation 2.122, one may add parts of the phenomenological model. For instance, the assumption of perfect mixing creates a link between the outlet streams and the inventories, therefore limiting the problems generated by the difficulties measuring inventories.

The various models and sub-models presented above must now be coupled with measurement values to subsequently perform data reconciliation. The next section describes some of the options for on-line and dynamic data reconciliation.

2.10.4 Reconciliation Methods

The state variables are usually not directly measurable. As in the steady-state and stationary methods, two main options can be used: either reconstructing pseudo-measurements of state variables by combining primary measured variables, or using the actually measured variables and expressing them as functions of the state variables. The second approach has the drawback of generating non-linear observation equations but the merit of using directly the raw information. As usual the measurement equation is

$$Y(t) = g[X(t)] + e \quad e \sim N(0, V), \quad (2.125)$$

where e is the measurement error and X the state variables, either f_i and m_i in the linear state equation case, or f_0 and c_i , total mass flowrates and component concentrations, in the bilinear case.

A variety of reconciliation methods are possible depending on the model and criterion used in the filter. The process equations can be those of a full model or a sub-model, empirical or phenomenological, steady-state, stationary, or fully dynamic. The reconciliation criterion usually contains weighted quadratic terms derived from the maximum likelihood estimation of the state variables X from the measurement Y , in a Gaussian context for model uncertainties, measurement errors, and driving white noise. The criterion can be instantaneous, or expressed in a finite time window, or launched at time zero. Depending on the consistency between the modeling assumptions and the criterion formulation and resolution, the resulting filter can be optimal or sub-optimal.

Let's now define the possible terms of a reconciliation criterion. They are written below at time t only, but can be extended in a finite size window, or a window beginning at time zero, thus usually leading to optimal observation. First, there are terms representing measurement errors, *i.e.* squared residuals according to

$$J_m(t) = [Y(t) - g[X(t)]]^T V^{-1} [Y(t) - g[X(t)]]. \quad (2.126)$$

Squared node imbalances can be put into the criterion when appropriate. They are

$$J_\varepsilon(t) = f(t)^T M^T V_\varepsilon^{-1}(0) M f(t). \quad (2.127)$$

When time correlation between the node imbalances ($V_\varepsilon(\tau) \neq 0$ for $\tau > 0$) is to be taken into account, the vector f is replaced by the stacked values of f at previous times, and the weighting matrix accordingly augmented with the autocovariance terms.

The model uncertainties δ and the driving white noise signals, either ξ , ξ' or ω , which are functions of the nature of the state variables to be estimated, should also be part of the reconciliation criteria and lead to the following quadratic terms:

$$J_\varsigma(t) = \varsigma(t)^T V_\varsigma^{-1} \varsigma(t), \quad (2.128)$$

where ς is either δ , ξ , ξ' , ω , depending upon the model used. Moreover, if a smoothing horizon is considered in the criterion formulation, past terms can also be used.

Table 2.6 summarizes some of the possibilities of coupling state equations, observation equations and quadratic reconciliation criteria. It gives the models and measurements that are used as well as the quadratic criteria. The methods are optimal or sub-optimal depending on the consistency between the modeling assumptions and the reconciliation criterion used. For instance, when the time correlation is quantified in the model while the criterion contains only present information (thus freezing past estimates) the method is sub-optimal. On the contrary, when the smoothing horizon is extended to all past information, the methods are optimal.

The criterion minimization may be a LQ optimization problem, but most frequently it is a non-linear optimization problem that requires using non-linear programming methods (see for instance [38]) or sequentially linearizing the observation equations. In the case where the estimation horizon starts at time zero and a full model is used, one could apply either the Kalman filter in the LQ case [99], or the extended Kalman filter (EKF) in the non-linear case. When a sub-model is used the generalized version of the Kalman filter (GKF) can be used [37, 96, 100].

It is also possible to use sub-models for the steady-state part of the model (thus ignoring the process gains) and a full model for its dynamic part [93]. This allows some kind of synchronization of the data, while avoiding constructing a complete model of the process. This technique has been used to deal with processes exhibiting large pure delays, as those occurring in pipe-lines [101].

There are also methods that keep the bilinear structure of the mass conservation equations (products of flowrates and concentrations), instead of using species flowrates as above, thus leading to linear observation equations. Estimation algo-

Table 2.6 Some of the possible dynamic reconciliation methods (h = smoothing horizon)

Method	Model	State equations	Criterion	Time window	References
On-line steady-state mass balance constraints	Sub-model	(2.123), $V_\varepsilon = 0$	(2.126)	$(t-h, t)$ or t	Makni and Hodouin, 1994 [92]
On-line stationary mass balance constraints (also called node imbalance method)	Sub-model	(2.123), $V_\varepsilon(0) \neq 0$	(2.126), (2.127)	$(t-h, t)$ or t	Bazin <i>et al.</i> 1995 [102], Hodouin and Makni 1998[24], Makni <i>et al.</i> 1995 [104], Lachance <i>et al.</i> 2006, 2007 [43, 76]
Idem with correlated node imbalances	Sub-model	(2.123), $V_\varepsilon(k) \neq 0$	(2.126), (2.127)	$(t-h, t)$ or t	Hodouin (unpublished)
Dynamic mass balance constraints	Sub-model	(2.122)	(2.126)	$(0, t)$	Almasy 1990 [36], Darouach and Zasadzinski 1991 [37], Lachance <i>et al.</i> 2006 [100]
Dynamic mass balance constraints plus empirical dynamics	Sub-model	(2.122), (2.124)	(2.126), (2.127), (2.128)	$(0, t)$	Almasy, 1990 [36], Lachance <i>et al.</i> 2006 [100]
Full phenomenological model	Full-model	(2.114) to (2.117)	(2.126), (2.127)	$(0, t)$	Liebman <i>et al.</i> 1992 [38]
Full empirical model	Full-model	(2.118) to (2.119)	(2.126), (2.128)	$(0, t)$	Bai <i>et al.</i> 2006 [39], Hodouin and Makni 1996 [93], Makni <i>et al.</i> 1995 [97]

rithms derived from the BILMATTM steady-state method [4, 92] can be used, which hierarchically decompose the optimization problem. The state variables are estimated in two steps. At the upper level the flowrates are estimated by minimizing the node imbalances calculated with the measured values of the concentrations. At the lower level, flowrate estimates are frozen, thus leading to a LQ problem, and the concentrations are reconciled. The procedure can be iterated between the two levels to improve the minimum localization. These techniques have been illustrated by numerical tests based on simulated or industrial data [46] and their drawbacks and advantages discussed [42].

The stationary method (node imbalance method) is recommended for industrial applications since it does not require any model parameter calibration, does not rely on inventory measurement, and allows reasonable adjustments of the dynamic (V_ε) and measurement error (V) variances (see, for instance, [43, 78], and [96]), which gives sufficient filtering performances to the method with respect to practical industrial needs.

2.10.5 An Example of Dynamic Reconciliation for a Simulated Flotation Circuit

A Pb–Cu–Zn flotation plant is depicted in Figure 2.16. It produces four different concentrates. On-stream analyzes (OSA) for lead, copper and zinc are available for the main streams of the plant (all the streams of Figure 2.16), but only the plant ore feedrate is measured. The plant is simulated using a phenomenological kinetic model [40] and random variations of the process variables are imposed as well as measurement errors. For data reconciliation a sub-model is considered that involves an empirical transfer function to describe the feed-to-concentrate and feed-to-reject dynamics for the ore and the three metals (only the ore flow model is given in Figure 2.17 but the flow models for the three metals have similar structures). However, the separation coefficients for the ore and the three strategic metals are not modeled. Consequently the model is incomplete.

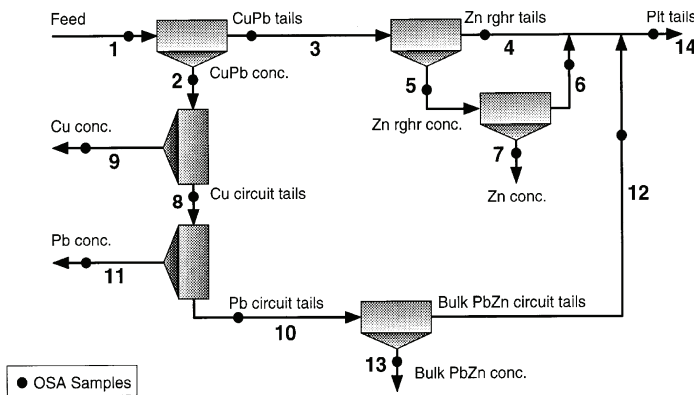


Figure 2.16 Flowsheet of a lead–zinc–copper flotation plant

The reconciliation procedure has a hierarchical structure (see Figure 2.18). At the upper level, metal assays are used to estimate flowrates in order to minimize model equation residuals at time t (this is a one-step method using criterion J of Equation 2.127). Assuming that ore flowrates are set at their estimated values, metal assays are reconciled at the lower level by minimizing squared residuals of the measurement equations and squared residuals of metal flowrate dynamic models. This is again a one-step method. There is also a possibility of refining flowrate estimation by iterating between the two levels, but this is usually unnecessary since the process variables do not vary much between two subsequent sampling times. Some results are given in Figures 2.19 to 2.22. The first one shows the estimated values of the unmeasured flowrates compared with the simulated values (“true values”). The other figures show the reconciled values of the metal assays for selected streams.

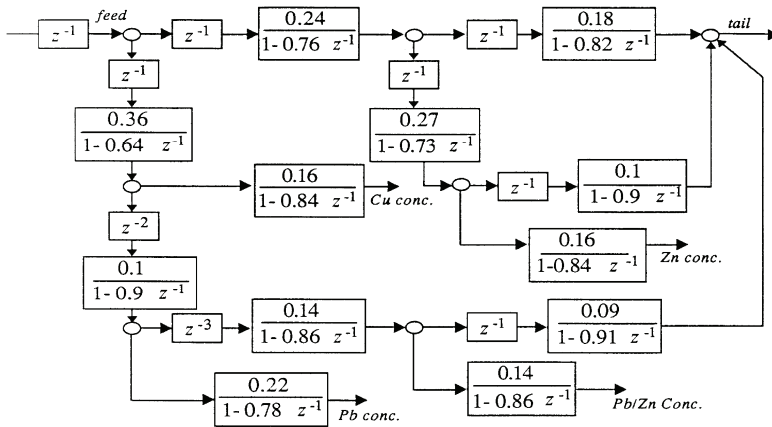


Figure 2.17 Empirical sub-model of the ore flow dynamics in the plant (z^{-1} in transfer functions is the backshift operator)

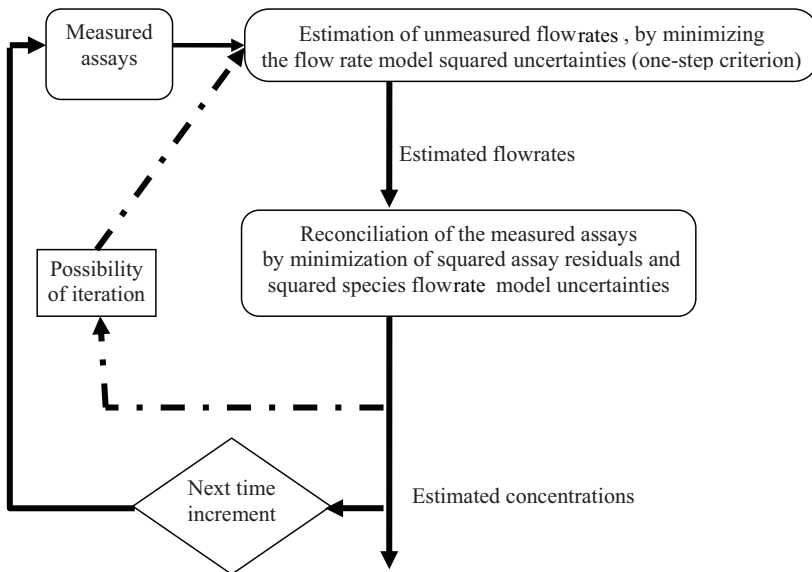


Figure 2.18 Computation scheme for reconciliation of Figure 2.17 plant data

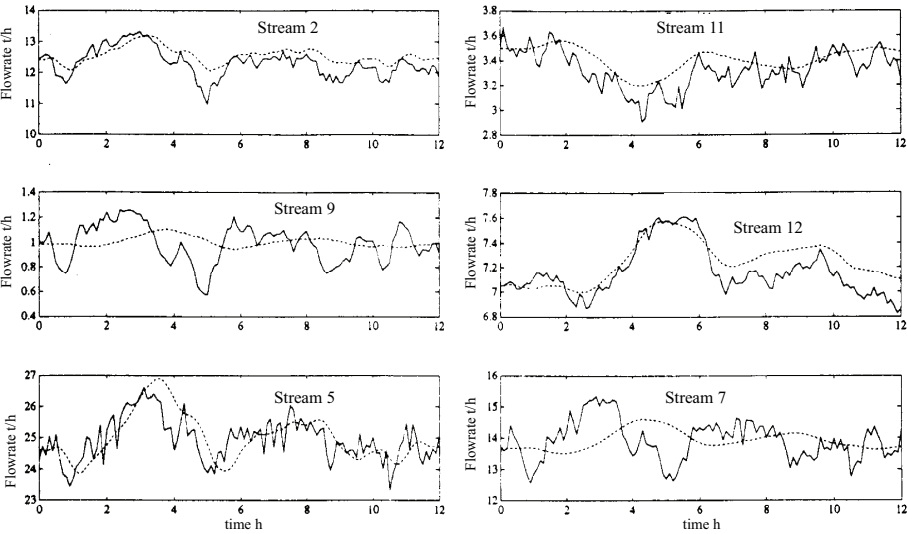


Figure 2.19 Comparison of estimated (dotted line) and simulated (continuous line) ore flowrates in six different streams

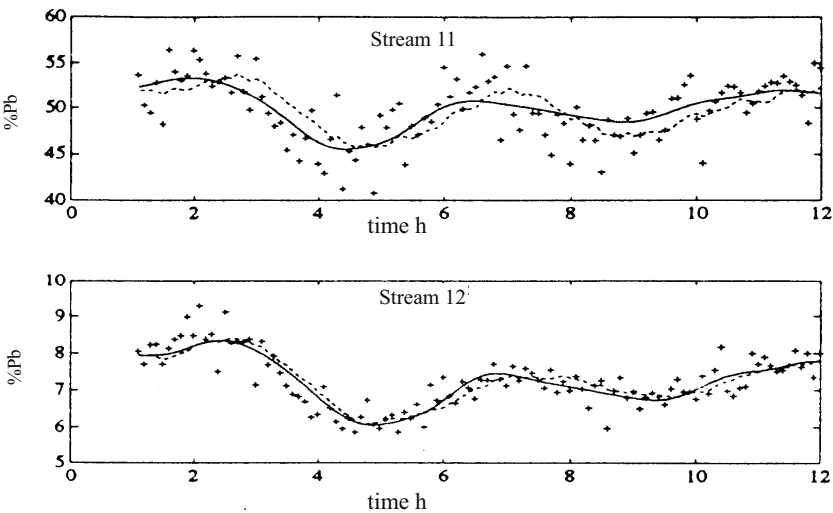


Figure 2.20 Comparison of measured (crosses), estimated (dotted line) and simulated (continuous line) lead grades in two selected streams

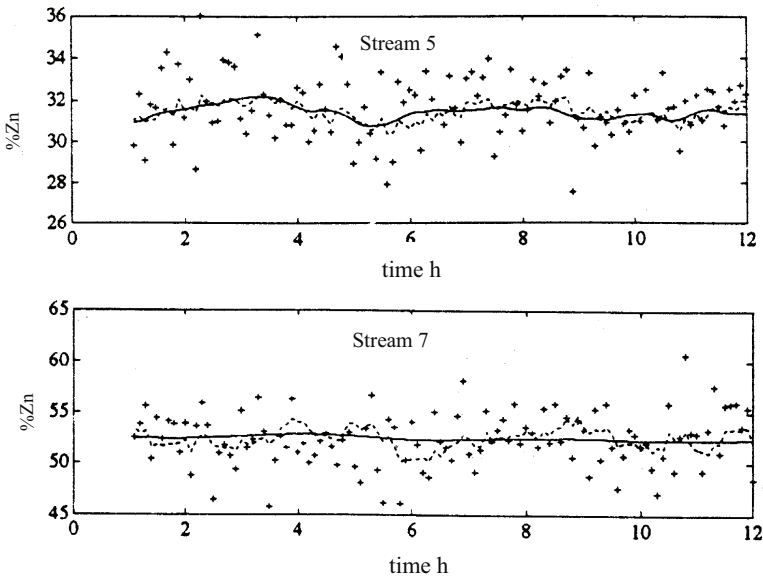


Figure 2.21 Comparison of measured (crosses), estimated (dotted line) and simulated (continuous line) zinc grades in two selected streams

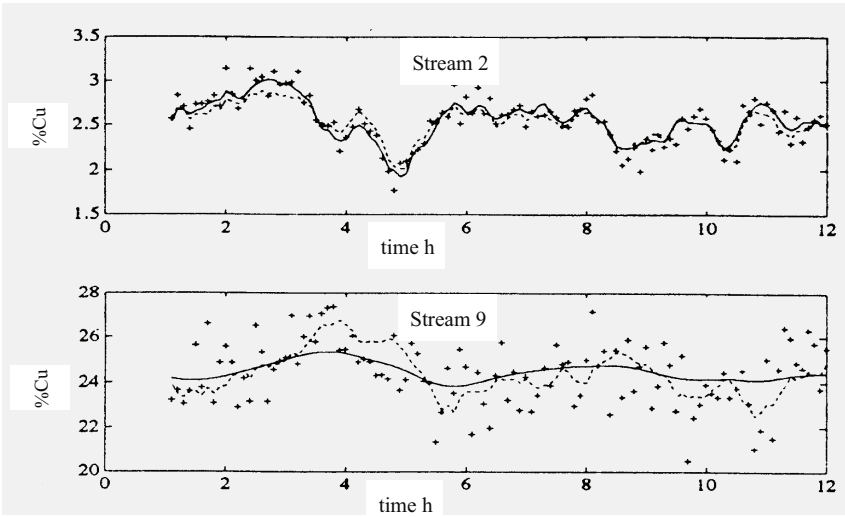


Figure 2.22 Comparison of measured (crosses), estimated (dotted line) and simulated (continuous line) copper grades in two selected streams

2.11 Instrumentation Strategy Design

If raw data delivered by sensors is to be updated by a reconciliation procedure, which at the same time generates estimates of unmeasured variables, the sensors, or more generally the variables to be measured, must be selected in a way to maximize the reconciliation procedure performance [51, 105–107]. The instrumentation placement or measurement strategy design problem is defined as [64, 108–110]:

Knowing the state equation $f(X) = \varepsilon$, or $CX = \varepsilon$, and V_ε (usually $= 0$ at the instrumentation design stage), find $g(\cdot)$, or C , and V_e in such a way that the following properties of the data reconciliation procedure are obtained, subject to the constraints that states X are observable:

- o *maximum accuracy of the estimated states and of the process performance indices that are subsequently calculated with these estimates;*
- o *Minimum instrumentation cost (capital investment, labor and maintenance costs);*
- o *maximum reliability of the process observer (maximum operating life without loss of observability, in the presence of possible total failures of sensors).*

This is a multi-criterion discrete optimization problem since the properties of the variables to be measured (type of sensor, accuracy of sensor, and sensor position in the process network) have to be selected from a discrete set of available instrumentation combinations. It is obvious that there is a necessary trade-off between the cost of the sensors and the reliability and accuracy of the observer, since increasing the number and accuracy of the sensors would improve both the observer accuracy and reliability. The instrumentation design can be processed either as a multi-objective problem or as a single-objective problem if the economical impact of improving the process operation control by an accurate and reliable observer can be quantified in the same units as the investment and maintenance costs. In the most general situation, there is no other available systematic method to find the best solution to the design problem than to scan the set of possible combinations. However, for specific cases such as the minimal cost combination or the minimal estimate variance for the minimal number of sensors, the optimal design can be found analytically, or the solution space scanned along paths minimizing the number of tested designs [67].

Table 2.7 illustrates the optimal instrumentation problem for the flotation plant of Figure 2.7 [52], using two different criteria (the cost and the estimate accuracy) and different allowed numbers of sensors for measuring the mineral flowrates.

Table 2.7 Optimal sensor placement for the plant of Figure 2.7 (numbers of the streams to be instrumented)

Number of available sensors	4	5	6
Minimal cost configuration	1, 6, 7, 8	1, 4, 6, 7, 8	1, 2, 4, 6, 7, 8
Minimal estimate variance configuration	5, 6, 7, 8	4, 5, 6, 7, 8	1, 4, 5, 6, 7, 8

2.12 Fault Diagnosis

The above data reconciliation procedures are unbiased insofar as the assumption that random components e and ε of the process model have zero means is valid. In the absence of such a property, the X estimates become biased and the data reconciliation, instead of improving the data and the subsequent control actions on the process, might distort quite significantly the process state observation and consequently deteriorate the decision making process involved in the manual or automatic control procedures [111].

Non-centered rates of accumulation may arise from omitted streams, corresponding either to infiltration of matter into the nodes, or, on the contrary, to material leakages, or to secondary inputs or intermittently active outputs, or to persistent deviations from the stationary operating conditions. Gross errors or accidental errors are also improbable events that may distort the reconciliation results [49, 112–116]. Either these faults should be first detected and corrected, or the reconciliation procedure robustified to attenuate their impacts on the reconciled values. Providing for such faults, in the linear data reconciliation case, Equation 2.21 can be more exactly rewritten as [70, 117]

$$MX = \varepsilon + L_F F_X, \quad (2.129)$$

where F_X is the vector of potential sources of non-centered accumulation rates, and L_F the matrix that distributes the impact of these faults on the various nodes. When L_F is the identity matrix, it is assumed that each node is potentially faulty and that all the potential faults around a node k are lumped into a single variable $F_X(k)$.

Non-centered measurements may arise from sensor biases due either to calibration deficiencies, matter sampling systematic errors, or to sensor misplacement in the streams to be measured. Providing for such faults, Equation 2.29 in the linear case can be more exactly rewritten as

$$Y = CX + e + K_F F_Y, \quad (2.130)$$

where K_F and F_Y have the same meanings for the measurement biases as L_F and F_X for the node balance faults. The fault detection problem can be defined as follows:

Knowing $M, C, Y, V, V_\varepsilon, L_F$, and K_F , decide between the two following hypotheses:

$H0 : F_X = F_Y = 0$, i.e., there are no faults in the process;

$H1 : \text{either } F_X \text{ or } F_Y \text{ or both are } \neq 0$, i.e., there is at least one fault in the process.

Most of the time, detection of gross errors is performed after the reconciliation procedure by testing innovations, i.e., corrections brought to the measured process variable values. When the statistical tests are sequentially applied to single innovations, these methods are incorrect since they ignore the fundamental correlation brought to the measured variable estimates by the reconciliation procedure. A gross error in any process variable usually contaminates the whole set of state estimates, and therefore might lead to wrong diagnosis. An alternative to the innovation residuals obtained through data reconciliation is to directly use the residuals of the redundancy equations. Various fault detection tests applied to the residuals of the redundancy equations (see Figure 2.5 and Equation 2.38) are available [7, 8]. Only the parity space approach is presented here. Each residual (element of the parity vector, [114]) can be tested against a normal distribution, or they can be tested together. The latter approach, called the global detection test, is applied to the following quadratic term:

$$J_r = r^T V_r^{-1} r, \quad (2.131)$$

where r represents the residuals of the redundancy equations (Figure 2.5 and Equation 2.38), hence a normal centred vector in the absence of faults since it depends only on e and ε . V_r , the variance matrix of the residuals, is directly calculable, in the linear case, since r consists of linear functions of e and ε (Equation 2.45). This quadratic term is a fault signature in the most usual case where the set of potential faults is not specified (no L_F and K_F matrices structuring the fault distribution). Since the observer uncertainties have been assumed to be normal, J_r follows a χ^2 statistical distribution with $m + q - n_X$ degrees of freedom and a non-central parameter only depending on F_X and F_Y , which are zero when there is no fault in the process. J_r is tested against a given level of false alarms.

When $H1$ is the conclusion of the fault detection step, the fault isolation step becomes [118]

Deciding between the two following hypotheses:

$H0 : \mathcal{F}_a = 0, \mathcal{F}_b \neq 0$,

$H1 : \mathcal{F}_a \neq 0, \mathcal{F}_b = 0$,

where $\mathcal{F}_a, \mathcal{F}_b$ are two exhaustive subsets of the \mathcal{F} set of faults ($F_X \cup F_Y$).

By repetitive application of this diagnosis test to various fault subsets, one can isolate the most probable active faults. The statistical isolation tests are performed on residuals derived from the redundant equations (Equation 2.45). The generalized

likelihood ratio test [118] is applied to the residuals of the assumed fault subsets, and the subset exhibiting the maximum score is assumed to be responsible for the detected faults [70, 117].

Example. Figure 2.23 gives an illustration of the method of fault isolation for the plant of Figure 2.7. A bias that has an amplitude of three times the measurement error standard deviation has been simulated in the sensor of stream 4 (in the data reconciliation example of Section 2.7.2, the stream 3 metal flowrate was assumed to be unmeasured; here all the states are measured, thus increasing the redundancy degree). The set of potential faults consists of the biases $F_Y(1)$ to $F_Y(8)$ for the eight streams and the node imbalances faults $F_X(1)$ to $F_X(4)$ for the four nodes. The diagnosis test is performed by sequentially splitting the fault set into one particular fault and the remaining potential faults. Knowing the statistical properties of e and ε as well as the fault amplitude, it is possible to calculate the probability of isolation of the fault. This is what has been done in Figure 2.23, assuming that a 5% level of false alarms is tolerated. It shows the probability of fault detection (D), while the other probability bars represent the probability of diagnosing the twelve potential faults as being the actual fault. It is clear that the assumption that the fault is in sensor 4 has the largest probability. However, wrong isolation decisions are also possible. The isolation test performance would obviously increase if the bias were of larger amplitude, or the data acquisition process repeated in a larger time window.

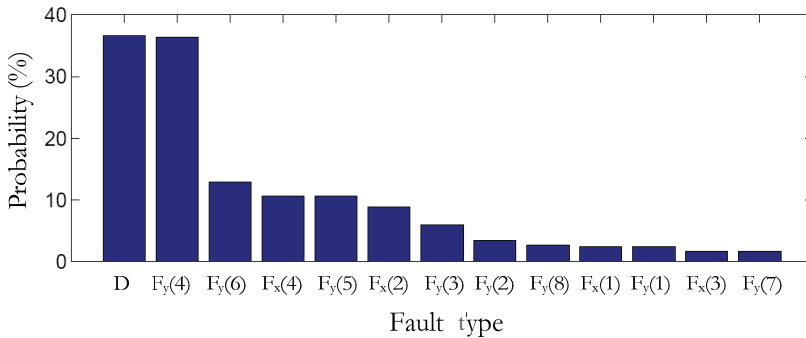


Figure 2.23 Probabilities of detection (D) and isolation of a bias in sensor 4, and of other potential faults F_X and F_Y

Figure 2.24 shows the same bar diagram for a simulated leakage at node number 2 of Figure 2.7. The amplitude of the leakage is three times the standard deviation of the node imbalance. Again the right diagnosis is the most probable, but wrong fault isolations are possible.

The same method can be applied to dynamic data reconciliation [116], [119, 120]) increasing the size of the vector Y by adding past state values, or by pre-filtering the measurements on a given past time horizon such that the node input and output measurements are synchronized [50].

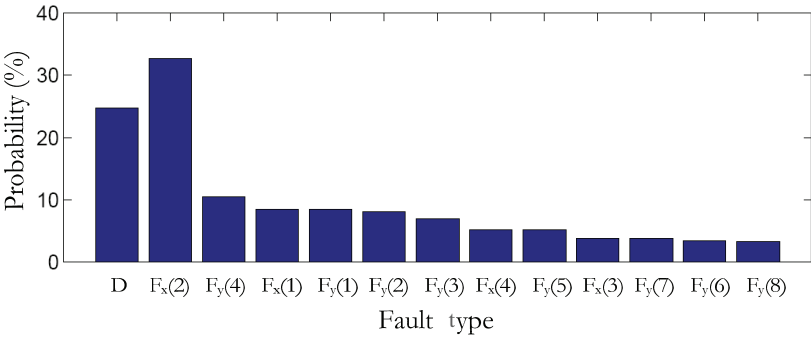


Figure 2.24 Probability of detection and isolation of a leakage at node 2

2.13 Coupling Data Reconciliation with Process Control and Optimization

As discussed in Chapter 2 introduction, data reconciliation has off-line applications such as process audit, metallurgical inventories, process modeling, plant design or reconfiguration, and tuning of operating conditions. Applications of on-line data reconciliation to automatic control and real-time optimization have also an interesting potential in process engineering, since, as in off-line applications, the use of better information would improve the performance of the methods that requires experimental data [47, 48, 121, 122]. Figure 2.25 gives a general scheme of the integration of a data reconciliation procedure into process control and optimization loops.

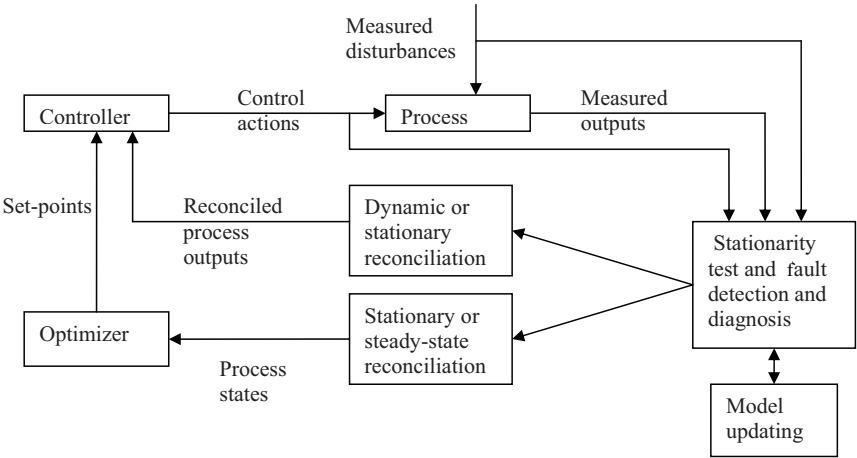


Figure 2.25 General scheme for the integration of data reconciliation techniques into automatic control and real-time optimization loops

Data reconciliation techniques are involved at two different levels when automatic control and real-time optimization are simultaneously implemented. The two observers must be different since the dynamics of the two loops are quite different, optimization being performed with slower dynamics than automatic control. This implies that the variances of the measurement errors and of the accumulation rates (node imbalances) must be tuned differently. Also, the models as well as the state variable of the two observers may be slightly different. Even if the same stationary reconciliation structure is used for both observers, the variances in the reconciliation criterion must be different. As the measured values in the optimization observer are obtained through an averaging technique involving a moving time window with several samples, their variances are usually divided by the number of samples in the window. Also, the accumulation rate variances are necessarily lower because the averaging process in the time window decreases the magnitude of the node imbalance variations, by attenuating the signal dynamics. Ultimately, the optimization observer could be steady-state.

In addition to the two observers, Figure 2.25 shows peripheral tools for data pre-processing. Sensor failures or abnormal process behavior must be detected before feeding the reconciled values to the optimizer or the controller. As the optimization observer is assumed working in stationary regime, it is important to test that the process variable means are statistically constant before reconciliation. Also when persistent mean changes are detected, it might be helpful to adapt model parameters, when, for instance, permanent changes occur to operating conditions such as ore grindability and grade, tonnage, chemical reagent type. If an adaptation procedure is integrated into the loops, it should be activated only when permanent changes due to persistent disturbance means or set-point changes are detected.

Flotation plant example: to illustrate the concept of control– optimization– reconciliation coupling, a simple example for a flotation plant is depicted in Figure 2.26. Data reconciliation is performed only at the optimization level and the control loop limited to a single-input-single-output system, where the collector addition is the manipulated variable and the concentrate grade the controlled variable. The grade set-point is supervised through the maximization of an economic index. Although there is no documented study of the performance of such a real-time optimization strategy, the concept has certainly a potential that should be investigated.

2.14 Conclusion

The objective of this chapter was to point out a problem that faces most metallurgical engineers and mineral processors who are involved with metal production and willing to understand and optimize the processes they are dealing with: the available measurements are uncertain, incomplete, and inconsistent with process behavior prior knowledge. The emphasis is put here on data reconciliation with mass conservation constraints, however this topic belongs to the universal problem of matching raw data and prior theoretical knowledge. The subject is superficially

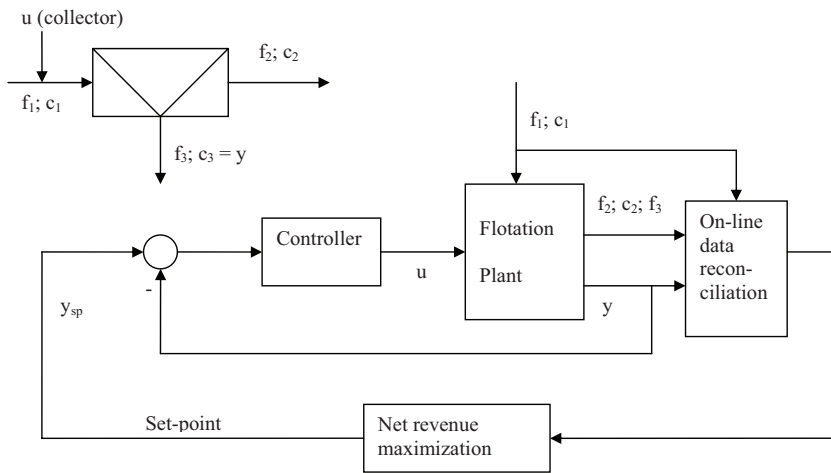


Figure 2.26 Real-time optimization of a flotation plant (f_j = ore flowrate; c_i = ore grade)

covered but the chapter explores techniques that could be helpful to practitioners as well as to researchers willing to deepen the concepts. The following issues arise from this limited presentation:

1. Measurement error sources are many and their variances are additives. There is a complex intricacy between the true measurement errors and the true dynamic variations of the process variables. It is most important to clarify these concepts at the beginning of a reconciliation procedure, to properly define the pursued reconciliation objectives.
2. The reconciliation objectives vary from the estimation of the underlying process steady-state operating regime to the fast tracking of its real instantaneous dynamic state. The method to be used for reconciliation must be adapted to the subsequent utilization of the reconciled data, which are mainly monitoring, modeling, control, and optimization;.
3. Assumptions that process variables variations and measurement errors are Gaussian and unbiased are frequently made. This is obviously not exactly true. This is why it is so important to make a prior detection of abnormal data or process behavior and a posterior analysis of the reconciliation residuals.
4. The reconciliation criterion may contain different types of residuals and weighting factors. Its formulation is essential to adequately match the reconciliation objectives, as well as the structure of data statistical properties. It must be rigorously designed, avoiding as much as possible users subjectivity or empirical tunings.
5. The reconciliation feasibility must be carefully investigated, by looking at redundancy of data and constraint information content.

6. As the success of data reconciliation methods strongly depends on the database structure, it can be used to design measurement strategies that promote estimate accuracy and reliability.

Nomenclature

The notation that is specific to illustrative examples in the text is not defined below.

Roman letters

A, B, A', B' : matrices of coefficients in a state space stochastic model

b : measurement bias

$B(\cdot)$: matrix function of the separation coefficients for steady-state modeling of a separation plant

c : composition factor

c_i and \hat{c}_i : concentration of component i and reconciled value

C : measurement coefficient matrix in the linear case

C_f : measurement coefficient matrix for the total mass flowrates

C_i : measurement coefficient matrix for the component i concentrations

C_i^d : diagonal matrix constructed with the vector c_i

d : sieve opening retaining 25% of the particles

d_{lib} : particle liberation size

D_i, E_i, F_i : matrices of coefficients associated to component i

e_I : integration error

e_r : measurement error of the redundant measured variables

e_f : measurement error of the total mass flowrates

$E(\cdot)$: mathematical expectation

$f(X)$: constraints of mass or energy conservation

f : vector of all the component flowrates (or shape factor in Section 2.4)

f_i : vector of component i flowrates in the plant network

$f_i(\cdot)$: i^{th} reconciliation constraint

f_0 : total mass flowrate vector

\bar{f}_0 : f_0 mean value

\bar{f} : f mean value

F_0^d : diagonal matrix of the vector f_0

F_X, F_Y : Faults in the conservation and measurements equations

$\mathcal{F}, \mathcal{F}_a, \mathcal{F}_b$: Union of the faults sets F_X and F_Y and two complementary subsets of it.

g : size distribution factor

$g(X)$: expression defining the measured variables

$G(z^{-1})$: discrete transfer function

h : time window width

$h(\cdot)$: function of the independent variables

h_i : component i mass fraction in node hold-ups

i : index of the steam components

in : innovations of the reconciliation procedure

I : identity matrix

$J(X)$: reconciliation criterion

J_m : part of the reconciliation criterion that contains measurements

J_ε : part of the reconciliation criterion that contains node imbalances
 J_ζ : part of the reconciliation criterion that contains driving white noise
 K_F : projection matrix of the faults in the measurements equations
 k : node index
 L : matrix for expressing state variables for independent state variables
 L_F : projection matrix of the faults in the conservation equations
 \mathcal{L} : Lagrangian function
 m : number of measured process variables
 m_i : vector of component i hold-up in the various nodes of the plant network
 m_0 : mass hold-ups of total material
 M_{dep} and M_{ind} : matrices associated to dependent and independent state variables
 M_i : network incidence matrix for component i
 M : incidence matrix of the plant network for all the stream components
 M_m and M_{um} : incidence matrices for measured and unmeasured state variables
 M_0 : incidence matrix for total mass
 M_s : sample mass
 n : number of components in plant streams (total mass not included)
 n_X : number of state variables
 n_n : number of plant network nodes
 p : number of streams in the plant network
 P_i : production rate of the component i
 q : number of conservation equations
 Q, Q', Q'' : functions or coefficient matrices in the deductible part of the constraint equations
 r : residuals of the redundancy equations (vector in parity space)
 R or $R(Y)$: coefficient matrix or function in the redundancy equations
 s : vector of node separation coefficient of a mineral separation plant
 s_{ki} : component i separation coefficient at node k
 $S(e)$: contributions of the measurement errors to the parity vector
 T or $T(\varepsilon)$: contributions of the constraint uncertainties to the parity vector
 $Var(.)$: statistical estimate of the variance of a process variable
 $V_{\hat{x}}$: variance of the reconciled states
 $V_x(k), V_x$: autocovariance of any process variable x and variance (*i.e.*, $V_x(0)$)
 V : variance of the measurement errors e
 $V_\varepsilon, V_{\varepsilon i}$: variances of ε and ε_i
 V_f : variance of f
 V_i : variance of c_i
 V_{in} : variance of the innovations
 V_r : variance of Y_r
 $V_{\hat{y}}$: variance of the reconciled measured variables
 V_δ, V_ω : variance of white noise $\delta(t)$ or $\omega(t)$
 X : state variable
 X_m : measured state variables
 X_{mr}, \hat{X}_{mr} : redundant measured state variables and reconciled value
 X_o, X_{or}, X_{onr} : observable state variables (redundant and non-redundant)

\hat{X}_o : reconciled value of X_o
 X_{no} : non-observable state variables
 X_{um} : unmeasured state variables
 x_i : state variable associated to component i
 x : any generic process variable (scalar or vector)
 x_s : sample mineral content
 x^* : true value of the mineral content of a sampled batch
 \hat{X} : reconciled state variable
 X_{dep} : dependent state variables
 X_{ind} : independent state variables
 \hat{X}_{ind} : reconciled independent state variables
 \hat{X}_{dep} : reconciled dependent state variables
 Y : measurement values
 Y_f : measurement values of f
 Y_i : measurement values of component i mass fractions
 Y_{nr} : measured values of the non-redundant measured variables
 Y_r : measured values of the redundant measured variables
 z : state space variables of a stochastic system
 z^{-1} : backshift operator
 Z : measured variables
 Z_{nr} : non-redundant measured variables
 Z_r : redundant measured variables

Greek letters

ε : overall conservation constraint uncertainties
 ε_0 : uncertainty of the total mass conservation equation
 ε_i : uncertainties in the conservation constraints of component i in the plant network
 γ_i : auxiliary matrix in the node imbalance solution
 $\delta(t)$: uncertainties in the process dynamic model
 λ : Lagrange's multipliers
 $\xi_i(t), \xi(t), \xi'(t)$: white noise
 $\xi_d(t)$: white noise driving uncertainties in the process dynamic model
 μ_x : mean value of x
 σ_e : standard deviation of the measurement error e
 σ_F : standard deviation of the fundamental sampling error
 $\sigma_{\hat{X}}$: standard deviation of the reconciled states
 Γ : auxiliary matrix in the linear reconciliation case
 ρ_v : degree of variance reduction
 $\rho(k)$: autocorrelation matrix
 ρ_{gan} : gangue density
 ρ_{min} : mineral density
 ρ_r : degree of redundancy
 Ψ and Φ : coefficient matrices in the linear case reconciliation information system
 ψ : matrix that extracts the node output streams from the vector f_i
 $\varphi(\cdot)$: function defining the node hold-ups

Π : auxiliary matrix in the expression of \hat{X}

τ : time lag for autocovariance definition

$\omega(t)$: white noise

Ω : matrix for the extraction of the feed stream flowrates from the plant stream flowrates

$\zeta(t)$: generic name for driving white noise

Acknowledgements This chapter is based on studies which have been undertaken by students and on ideas discussed with colleagues. The list is long but their names deserve to be given: Nicole Alliot, Gilles Barbéry, Claude Bazin, Steve Bellec, Antoine Berton, Yves Bérubé, Charles-Eudore Boudreault, Cameron Crowe, Luiz Rogerio Pinho de Andrade Lima, André Desbiens, Yan Guang Du, Marc-Denis Everell, Frédéric Flament, Thierry Gelpe, Luc Lachance, Chéfi Ketata, Daniel Laguitton, Sami Makni, El Hassan Mazzour, Jacques McMullen, Azar Mirabedini, Éric Poulin, Éric Plamondon, Simon Rochon-Tremblay, Jules Thibault, and Solange Vaz Coehlo.

References

- [1] Hodouin D, Jämsä-Jounela SL, Carvalho MT, Bergh L (2001) State of the art and challenges in mineral processing control. *Control Engineering Practice* 9(9):995–1005
- [2] Wiegel RL (1972) Advances in mineral processing material balances. *Canadian Metallurgical Quarterly* 11(2):413–424
- [3] Smith HW, Ichiyen N (1973) Computer adjustment of metallurgical balances. *CIM Bulletin* 66:97–100
- [4] Hodouin D, Everell MD (1980) A hierarchical procedure for adjustment and material balancing of mineral process data. *International Journal of Mineral Processing* 7(2):91–116
- [5] Kuehn DR, Davidson H (1961) Computer control II: Mathematics of control. *Chemical Engineering Progress* 57(6):44–47
- [6] Crowe CM (1996) Data reconciliation-progress and challenges, *Journal of Process Control* 6:89–98
- [7] Narasimhan S, Jordache C (1999), *Data reconciliation & gross error detection: an intelligent use of process data*. Gulf Pub. Co., Houston.
- [8] Romagnoli JA, Sanchez MC (2000) *Data processing and reconciliation for chemical process operations*. Academic Press
- [9] Algosys Inc. (2004) *Bilmat Technical Guide Version 9.0: An advance computer program for data reconciliation through mass balancing*, Quebec City, Canada
- [10] Caspeo (2008) www.caspeo.net
- [11] JKTech: www.jktech.com
- [12] Banisi: www.banisi.ir
- [13] OSIssoft: www.osissoft.com
- [14] IPS: www.ips.invensys.com
- [15] AspenTech: www.aspentech.com
- [16] Belsim: www.belsim.com
- [17] Du YG, Hodouin D, Thibault J (1997a) Use of a novel autoassociative neural network for nonlinear steady-state data reconciliation. *AIChE Journal* 43(7):1785–1796
- [18] Du YG, Thibault J, Hodouin D (1997b) Data reconciliation for simulated flotation process. *Artificial Intelligence in Engineering* 11(4):357–364
- [19] Aldrich C, Van Deventer J (1994) Identification of gross errors in material balance measurements by means of neural nets. *Chemical Engineering Science* 49(9):1357–1368

- [20] Meert K (1998) A real-time recurrent learning network structure for data reconciliation. *Artificial Intelligence in Engineering* 12(3):213–218
- [21] Mandel D, Abdollahzadeh A, Maquin D, Ragot J (1998) Data reconciliation by inequality balance equilibration: a LMI approach. *International Journal of Mineral Processing* 53(3):157–169
- [22] Hodouin D, Kasongo T, Kouamé É, Everell MD (1981) BILMAT: an algorithm for material balancing mineral processing circuits: applications to comminution, desliming and flotation units. *CIM Bulletin* 74:123–131
- [23] Bazin C, Hodouin D (1996) Processing assays of size fractions from sieve and cyclosizer analyses. *Minerals Engineering* 9(7):753–763
- [24] Hodouin D, Makni S (1998) Data reconciliation for stationary multi-phase multi-stream metallurgical processes. *Proceedings of the IFAC Automation in Mining, Mineral and Metal Processing, Köln, Deutschland*, 49–54
- [25] Laplante AR (1984) Plant sampling and mass balancing for gold ores. *Proceedings of the 1st Int. Symp. on Precious Metals Recovery*, 1–25
- [26] De Andrade LRP (2006), Nonlinear data reconciliation in gold processing plant. *Minerals Engineering* 19:938–951
- [27] Cimon D, Barbery G, Flament F, Hodouin D (1987) Materials balance in gold processing plants. *Proceedings Gold Metal. Int. Symp.*, 9–21
- [28] Bellec S, Hodouin D, Bazin C, Duchesne C (2007) Multi-level data reconciliation - application to a gold ore processing plant. *Proceedings of IFAC MMM07 Symposium*, 39–44
- [29] Bazin C, El-Ouassiti K, Hodouin D, Zouadi M (2005a) Data reconciliation in hydrometallurgy: applications to leaching of clay and copper solvent extraction. *Proceedings Conference of Metallurgists, Calgary, ed. METSOC CIM*
- [30] Bazin C, Hodouin D, Zouadi M (2005b) Data reconciliation and equilibrium constant estimation: Application to copper solvent extraction. *Hydrometallurgy* 80(1):43–53
- [31] Eksteen JJ, Frank SJ, Reuter MA (2002) Dynamic structures in variance based data reconciliation adjustments for a chromite smelting furnace. *Minerals Engineering* 15:931–943
- [32] Bazin C, Hodouin D, Duchesne C, Thibault J, Trusiak AR (1998) Reconciliation of mass and energy data measurements: application to a rotary dryer. *Canadian Metallurgical Quarterly* 37(3-4):333–342
- [33] Bazin C, Rochon-Tremblay S, Gosselin C (2003) Estimation of gas flow rates and pellets temperature in an iron oxide induration furnace. *Canadian Metallurgical Quarterly* 42:301–312
- [34] Hodouin D, Gelpe T, Everell MD (1982) Sensitivity analysis of material balance calculations - an application to a cement clinker grinding circuit. *Powder Technology* 3:139–153
- [35] Maquin D, Adrot O, Ragot J (2000) Data reconciliation with uncertain models. *ISA Transactions* 39:35–45
- [36] Alması GA (1990) Principles of dynamic balancing. *AIChE Journal* 36:1321–1330
- [37] Darouach M, Zasadzinski M (1991) Data reconciliation in generalized linear dynamic systems. *AIChE Journal* 37(2):193–201
- [38] Liebman MJ, Edgar TF, Lasdon LS (1992) Efficient data reconciliation and estimation for dynamic processes using nonlinear programming techniques. *Computers and Chemical Engineering* 16:963–986
- [39] Bai S, Thibault J, McLean DD (2006) Dynamic data reconciliation: alternative to Kalman filter. *Journal of Process Control* 16:485–498
- [40] Hodouin D, Lachance L, Desbiens A (2007) Dynamic Data Reconciliation: From the full model observer to the stationary on-line mass conservation filter - the flotation case. *Proceedings of IFAC Symposium on Automation in Mining, Mineral and Metal Processing, Québec, Canada*, 87–92
- [41] Darouach M, Ragot J, Zasadzinski M, Krzakala G (1989) Maximum likelihood estimator of measurement error variances in data reconciliation. *IFAC AIPAC Symposium, Nancy, France*, 109–112.

- [42] Hodouin D, Bazin C, Makni S (1993) On-line reconciliation of mineral processing data. Proceedings of the AIME/SME Symposium Emerging Computer Techniques for the Mineral Industry. Reno, Nevada, 101–110
- [43] Lachance L, Poulin É, Hodouin D, Desbiens A (2007) Tuning stationary observers: Application to a flotation unit simulator. Proc. of IFAC MMM Automation, Qubec City, Canada.
- [44] Hodouin D, Flament, Bazin C (1989) Reliability of material balance calculations - a sensitivity approach. Minerals Engineering 2(2):157–170
- [45] Bazin C, Franklin M (1996) Real-time material balance for flotation plants using a least-squares recursive algorithm. International Journal of Mineral Processing 46(3–4):231–244
- [46] Hodouin D, Bazin C, Makni S (1997) Dynamic material-balance algorithms: application to industrial flotation circuits. Minerals and Metallurgical Processing 14(2):21–28
- [47] Abu-el-zeet ZH, Roberts PD, Becerra VM (2002). Enhancing model predictive control using dynamic data reconciliation. AIChE Journal 48:324–332
- [48] Ramamurthi Y, Sistu PB, Bequette BW (1993) Control-relevant dynamic data reconciliation and parameter estimation. Computers and Chemical Engineering 17(1):41–59
- [49] Tong H, Crowe CM (1997) Detecting persistent gross errors by sequential analysis of principal components. AIChE Journal 43(5):1242–1249
- [50] Berton A, Hodouin D (2007) Synchronized node imbalances for fault detection and isolation in plant networks involving material recirculation. Computers and Chemical Engineering 31:815–832
- [51] Bagajewicz M (2001) Process plant instrumentation design and upgrade. Technomic Publishing Company
- [52] Mazzour EH, Hodouin D, Makni S (2002) Optimal sensor implementation in metallurgical plants - an application to a generic mineral separation plant. International Journal of Mineral Processing 69:185–203
- [53] Hodouin D, Vaz Coelho S (1987) Mass balance calculations around mineral processing units using composition analyses within particle-size classes. International Journal of Mineral Processing 21:65–82
- [54] Bellec S, Jiang T, Kerr B, Diamond M, Stuart P (2007) On-line processing and steady-state data reconciliation of pulp and paper mill process data. Pulp & Paper Canada 108:36–40
- [55] Bazin C, Hodouin D (2001) Importance of covariance in mass balancing of particle size distribution data. Minerals Engineering 14(8):851–860
- [56] Gy P (1979) Sampling of particulate materials. Elsevier, Amsterdam.
- [57] Pitard (1992) Pierre Gy's sampling theory and sampling practice. CRC Press, Baton Rouge, LA, Vols 1 and 2
- [58] Bazin C, Hodouin D, Grant R (1996) Semi-batch flotation tests: reliability of performance indices. Canadian Metallurgical Quarterly 35(4):321–327
- [59] Hodouin D, Garon M, Rémillard M, Thérien M (1988) Assessment of precious metal distribution in Lac Mattagami flotation plant by computer mass balance calculation. CIM Bulletin 81:62–69
- [60] Alhaj-Dibo M, Maquin D, Ragot J (2008) Data reconciliation: a robust approach using a contaminated distribution. Control Engineering Practice 16(2):159–170
- [61] Karjala TW, Himmelblau DM (1994) Dynamic data rectification by recurrent neural networks vs. traditional methods. AIChE Journal 40(11):1865–1875
- [62] Kretsovalis A, Mah RSH (1988) Observability and redundancy classification in generalized process networks. Computer and Chemical Engineering 16:689–703
- [63] Ragot J, Luong M, Maquin D (1996) Observability of systems involving flow circulation. International Journal of Mineral Processing 47:125–140
- [64] Hodouin D, Mazzour EH (2006) Robustness of metallurgical process observers based on mass and energy conservation constraints. Proceedings of a Workshop on Automation in Mining, Mineral and Metal Processing, Krakow, Poland, 167–172
- [65] Vaclavcek V (1974) How long to take the balance periods at balance calculations of quasi-stationary continuous processes. Chemical Engineering Science 29:2307–2313

- [66] Vaclavek V, Loucka M (1976) Selection of measurements necessary to achieve multicomponent mass balance in chemical plants. *Chemical Engineering Science* 31:1199–1205
- [67] Mazzour EH, Hodouin D, Makni S (2003) Une aide algorithmique à l'optimisation du placement des capteurs dans un procédé. *Journal Européen des systèmes automatisés* 37:1251–1276
- [68] Crowe CM, Garcia Campos GY, Hrymak A (1983) Reconciliation of process flow rates by matrix projection. *AIChE Journal* 29: 881–888
- [69] Kelly JD (1998) On finding the matrix projection in the data reconciliation solution. *Computers and Chemical Engineering* 22(11):1553–1557
- [70] Berton A, Hodouin D (2002) Linear and bilinear fault detection and diagnosis based on mass and energy balance equations. *Control and Engineering Practice* 11:103–113
- [71] Sanchez M, Romagnoli JA (1996) Use of orthogonal transformations in data classification-reconciliation. *Computers and Chemical Engineering* 20(5):483–493
- [72] Hodouin D, Alliot N, Flament F (1991) Redundancy analysis of complex sets of mineral processing data for mass balance computation. *International Journal of Mineral Processing* 32(3–4):213–231
- [73] Almasy GA, Mah RSH (1984) Estimation of measurement error variances from process data. *Industrial Engineering Chemistry Process Research and Development*. 23:779–784
- [74] Keller JY, Zasadzinski M, Darouach M (1992). Analytical estimator of measurement error variances in data reconciliation. *Computers and Chemical Engineering* 16:185–188
- [75] Chen J, Bandoni A, Romagnoli JA (1997) Robust estimation of measurement error variance/covariance from process sampling data. *Computers and Chemical Engineering* 21:593–600
- [76] Lachance L, Poulin É, Hodouin D, Desbiens A (2006b) Performance of steady-state and stationary data reconciliation as a function of process disturbances dynamics. *Mineral Process Modeling, Simulation and Control Conference*, Laurentian University, Sudbury, Canada, 393–408
- [77] Poulin É, Hodouin D, Lachance L (2009) Impact of plant dynamics on the performance of steady-state data reconciliation. submitted to *Computers and Chemical Engineering*
- [78] Poulin É, Hodouin D, Lachance L (2009) Estimation of measurement error variances in data reconciliation using a flow distribution model. *Proc. IFAC SYSID'2009*, St Malo, France
- [79] Hodouin D, Ketata C (1993) Variance of average stream compositions obtained by automatic incremental sampling. *International Journal of Mineral Processing* 40(3–4):199–223
- [80] Mirabedini A, Hodouin D (1998) Calculation of variance and covariance of sampling errors in complex mineral processing systems, using state-space dynamic models. *International Journal of Mineral Processing* 55(1):1–20
- [81] Maquin D, Ragot J, Darouach M, Fayolle J (1988) Validation des mesures par équilibre hiérarchisé de bilans-matière. *International Journal of Mineral Processing* 23:241–252
- [82] Schraa OJ, Crowe CM (1996) The numerical solution of bilinear problems using unconstrained optimization methods. *Event of the European Federation of Chemical Engineers* 20:S727–S730.
- [83] Stanley GM, Mah RSH (1977) Estimation of flows and temperatures in process networks. *AIChE Journal* 23(5):642–650
- [84] Kelly JD (2004) Techniques for solving industrial nonlinear data reconciliation problems. *Computers and Chemical Engineering* 28:2837–2843
- [85] Arora N, Biegler LT, Heyen G (2002) Data reconciliation framework. Software architecture and tools for computer aided process engineering. *Computer Aided Chemical Engineering* 11:193–212
- [86] Schladt M, Hu B (2007) Soft sensors based on nonlinear steady-state data reconciliation in the process industry. *Chemical Engineering and Processing* 46(11):1107–1115
- [87] Bazin C (1999) Personal communication.
- [88] Flament F, Hodouin D, Bazin C (1986) Propagation of measurement errors in mass balance calculation of mineral processing data. *Proceedings of APCOM'86*, Pennsylvania State University Editor, SME-AIME.

- [89] Hodouin D, Flament F (1991) Influence of data collection and conditioning strategies on the significance of performance indices in mineral processing plants. *Proceedings Int. Symp. Evaluation and Optimization of Metallurgical Performance, SME/AIME*, 195–208
- [90] Hodouin D, Bazin C, Trusiak A (1984) Reliability of calculation of mineral process efficiencies and rate parameters from balanced data. *Proceedings of Control 84, JA Herbst Editor, SME-AIME, Chapter 16*, 133–144
- [91] Reimers C, Werther J, Gruhn G (2008) Flowsheet simulation of solids processes. data reconciliation and adjustment of model Parameters. *Chemical Engineering and Processing* 47:138–158
- [92] Makni S, Hodouin D (1994) Recursive BILMAT algorithm: an on-line extension of data reconciliation techniques for steady-state bilinear material balance. *Mineral Engineering* 7(9):1179–1191
- [93] Hodouin D, Makni S (1996) Real-time reconciliation of mineral processing plant data using bilinear material balance equations coupled to empirical dynamic models. *International Journal of Mineral Processing* 48:245–264
- [94] Hlavacek V (1977) Analysis of a complex plant steady state and transient behavior. *Computers and Chemical Engineering* 1:75–100
- [95] Fillon M, Meyer M, Pingaud H, Enjalbert M (1996) Efficient formulation for batch reactor data reconciliation. *Industrial and Engineering Chemistry Research* 35(7):2288–2298
- [96] Lachance L, (2007) Observation de procédés basée sur des sous-modèles - Applications au traitement et au transport de la matière. PhD. thesis, Université Laval, Québec City, Canada
- [97] Makni S, Hodouin D, Bazin C (1995) On-line data reconciliation by minimization of a weighted sum of squared residuals and node imbalances. *Proceedings of the XIX International Mineral Processing Conference, San Francisco, USA*
- [98] Dochain D (2003) State and parameter estimation in chemical and biochemical processes: a tutorial. *Journal of Process Control* 13: 801–818
- [99] Makni S (1990) Le filtrage Kalman des ateliers de flottation. Master thesis, Université Laval, Québec, Canada
- [100] Lachance L, Hodouin D, Desbiens A (2006a) Using sub-models for dynamic data reconciliation. *Proceedings of IFAC ADCHEM Int. Symp., Gramado, Brazil*, 711–716
- [101] Lachance L, Hodouin D, Desbiens A, Poulin É (2008) Data reconciliation through mass balancing for flowsheets involving plug-flow units. *International Mineral Processing Congress, Beijing, China*
- [102] Bazin C, Trusiak AR, Hodouin D (1995) Application of a dynamic material balance program to an industrial flotation circuit. *CIM Bulletin*. 88:47–54
- [103] Hodouin D, Mirabedini A, Makni S, Bazin C (1998) Reconciliation of mineral processing data containing correlated measurement errors. *International Journal of Mineral Processing* 32:201–215
- [104] Makni S, Hodouin D, Bazin C (1995b) A recursive node imbalance method incorporating a model of flowrate dynamics for on-line material balance of complex flowsheets. *Minerals Engineering* 8(7):753–766
- [105] Ali Y, Narasimhan S (1995) Redundant sensor network design for linear processes. *AIChE Journal* 41(10):2237–2249
- [106] Bagajewicz M, Sanchez M (2000a) Reallocation and upgrade of instrumentation in process plants. *Computers and Chemical Engineering* 24(8):1945–1959
- [107] Bagajewicz M, Sanchez M (2000b) Cost-optimal design of reliable sensor networks. *Computers and Chemical Engineering* 23(11-12):1757–1762
- [108] Hodouin D, Mazzour EH (2004) Impact of sensor implementation strategies on the efficiency of data reconciliation. *Proceedings of the IFAC Symposium on Mineral, Metal and Materials Processing Automation, Nancy, France*
- [109] Hodouin D, Berton A, Mazzour EH (2004) Sensor placement, fault detection and data reconciliation using heat and mass conservation constraints. *Proceedings of Control 2004, PAP-TAC, Québec City*, 303–306

- [110] Hodouin D, Mazzour EH (2008) Measurement accuracy selection for designing observers of metallurgical plant performances. Proceedings of MCA-IEEE MED'08 Symposium, Ajaccio, France
- [111] Özyurt DB, Pike RW (2004) Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes. *Computers and Chemical Engineering* 28:381–402
- [112] Romagnoli JA, Stephanopoulos G (1981) Rectification of process measurement data in the presence of gross errors. *Chemical Engineering Science* 36:1849–1863
- [113] Tamhane AC, Mah RSH (1985) Data reconciliation and gross error detection in chemical process networks. *Technometrics* 27:409–422
- [114] Ragot J, Aitouche A, Kratz F, Maquin D (1991) Detection and location of gross errors in instruments using parity space technique. *International Journal of Mineral Processing* 31:281–299
- [115] Bagajewicz M, Jiang Q (1998) Gross error modeling and detection in plant linear dynamic reconciliation. *Computers and Chemical Engineering* 22(12):1789–1809
- [116] Chen J, Romagnoli JA (1998) A strategy for simultaneous dynamic data reconciliation and outlier detection, *Computers and Chemical Engineering* 22(4-5):559–562
- [117] Berton A, Hodouin D, (2004) A FDI procedure based on linearized bilinear mass and heat conservation models. Proceedings of the IFAC Symposium on Mineral, Metal and Materials Processing Automation, Nancy, France
- [118] Basseville M (1997) Information criteria for residual generation and fault detection and isolation. *Automatica* 33:783–803
- [119] Albuquerque JS, Biegler LT (1996) Data reconciliation and gross-error detection for dynamic systems. *AIChE Journal* 42(10):2841–2856
- [120] Bagajewicz M, Jiang Q (2000) Comparison of Steady State and Integral Dynamic Data Reconciliation. *Computers and Chemical Engineering* 24:2367–2383
- [121] Bai S, McLean DD, Thibault J (2005) Enhancing controller performance via dynamic data reconciliation. *Canadian Journal of Chemical Engineering* 83:515–526
- [122] Bai S, McLean DD, Thibault J (2007) Impact of model structure on the performance of dynamic data reconciliation. *Computers and Chemical Engineering* 31:127–135

Advanced Control and Supervision of Mineral
Processing Plants

Sbárbaro, D.; Del Villar, R. (Eds.)

2010, XX, 312 p. With online files/update., Hardcover

ISBN: 978-1-84996-105-9