Use of a Novel Autoassociative Neural Network for Nonlinear Steady-State Data Reconciliation

Yang-Guang Du and Daniel Hodouin
Dept. of Mining and Metallurgy, Laval University, Sainte-Foy, P.Q., Canada G1K 7P4
Jules Thibault
Dept. of Chemical Engineering, Laval University, Sainte-Foy, P.Q., Canada G1K 7P4

A novel autoassociative neural-network-based estimator for nonlinear steady-state data reconciliation was developed, which is a modified autoassociative feedforward neural network. The main difference between them lies in the minimization of an objective function that includes material imbalance terms of flow rates and compositions as well as the traditional least-square prediction term. Accordingly, this neural network, with the material balance-related equations included in the objective criterion, can perform simultaneously the following basic functions necessary for proper steady-state data rectification: (1) eliminate the nonrandom errors, such as the biases and gross errors, from measurements; (2) filter out the random errors from measured data; and (3) estimate the values of unmeasured process variables, provided data redundancy prevails. This novel neural-network-based reconciliation method is demonstrated on a flotation system.

Introduction

The last decades have witnessed the implementation of distributed control systems and a tremendous growth in the level of instrumentation in large chemical and mineral processing industries that now have on-line access to massive data banks. The key process variables are normally used for closed-loop control, whereas the great majority of measured variables remain unused or merely serve for monitoring the states of the process. A lot of research is now being conducted to find ways to make better use of this wealth of information that characterizes the real behavior of the process: development of softsensors (Aynsley et al., 1993), fault diagnosis (Dunia et al., 1996; Hoskins et al., 1991; MacGregor, 1995; Tong and Crowe, 1996; Wise and Ricker, 1991), statistical process monitoring (Chen et al., 1996; Nomikos et al., 1994; Piovoso and Kosanovich, 1994), and data reconciliation (Crowe et al., 1983; Crowe, 1986, 1996; Hodouin and Everell, 1980; Karjala and Himmelblau, 1994; Liebman et al., 1992; Mah, 1990; Romagnoli and Stephanopoulos, 1981; Schraa and Crowe, 1996; Smith and Ichiyen, 1973). The latter topic is the subject of this article.

Measurements of process variables are generally contaminated with significant random and nonrandom errors. For instance, in the flotation circuit used in this investigation, key process variables such as flow rates and stream mineral compositions are usually measured on-line using magnetic flowmeters and gamma-ray densitometers. Because of measuring difficulties related to the three-phase system, along with the complex measuring procedures and calibration problems, the measurement errors of these variables can be appreciable. The same is true for many other processes encountered in the chemical and mineral processing industries. To improve the level of confidence in these measurements and to rely on more accurate data for better process operation and control, it is important that these measurements be corrected, using some type of data-reconciliation method. Data reconciliation is a computing procedure that is currently used to adjust measured process variables to render them as coherent as possible to their true values, by forcing them to be more consistent with known process constraints such as material and energy balances (Schraa and Crowe, 1996).

The adjustment of process data, leading to better estimates of the true values of the process variables, is normally performed in two steps. Nonrandom measurement errors, such as persistent gross errors, must first be detected and then removed or corrected. Indeed, meaningful data adjustments can be achieved if and only if there is no gross error present

Correspondence concerning this article should be addressed to J. Thibault.

in the data (Tong and Crowe, 1996). Gross-error detection has been the subject of a significant amount of research (Mah, 1990; Rollins and Davis, 1992; Tjoa and Biegler, 1991; Tong and Crowe, 1996). The second step, usually referred to as data reconciliation *per se*, generally satisfies the following three objectives: (1) the random errors in measurement are filtered out in order to reduce the variance of the data; (2) the measured data are corrected to render them as coherent as reasonable with the process model and/or process constraints; (3) the values of unmeasured process variables are estimated. It is important to mention that data reconciliation is possible if and only if data redundancy is present in the data (Crowe, 1989; Zasadzinski, 1990).

By and large, all data-reconciliation methods essentially use a very similar approach. All the methods use both the available measurements and some prior knowledge, expressed in terms of process constraints such as mass and energy balances and/or more sophisticated static or dynamic models. Since neither perfect measurements nor perfect process models exist, reconciliation of data is performed by minimizing an objective function in which weighting coefficients are used to give the relative confidence level of each contributing term. Data reconciliation of steady-state processes has been the subject of significantly more research than it has for dynamic processes.

For steady-state processes, the problem usually boils down to the minimization of a least-squares objective function subject to some material balance constraints. Very often, the problem to be solved is bilinear. This problem has been studied extensively as in its two limiting cases: (1) an exact material conservation model, that is, all the confidence is placed on the model (Wiegel, 1972; Hodouin and Everell, 1980; Mah, 1990; Crowe, 1989; Meyer et al., 1993), and (2) exact measurements (Lynch, 1977). A two-step approach, where the bilinear system is solved in two successive steps in order to simplify the problem, has been proposed by Hodouin et al. (1982). More recently a method that balances both model and measurement uncertainties for steady-state and dynamic data reconciliation has been proposed (Makni et al., 1995). Accordingly, data reconciliation becomes an unconstrained minimization or a penalty function optimization problem. Nonlinear programming technique appears to be very well suited for the solution of the bilinear problem (Smith and Ichiyen, 1973; Liebman et al., 1992). A data-reconciliation method, called matrix projection, has been proposed by Crowe et al. (1983) for linear systems, whereby a reduced set of balance equations is obtained by eliminating the unmeasured variables. Crowe (1986) has extended the method of matrix projection to the bilinear systems. Schraa and Crowe (1996) have recently proposed solving the same bilinear problem, using an unconstrained optimization method based on analytical derivatives, but the performance was not as good as expected. Other approaches with filtering characteristics have also been proposed, such as the evaluation of reconciled values of process variables in a moving data window (Hodouin et al., 1993; Makni and Hodouin, 1994).

Data reconciliation for dynamic systems is significantly more complex because the process model must initially be derived or calibrated from raw measurements, that is, with process data not yet reconciled. In such systems, the functions of system identification and data reconciliation must be performed simultaneously, with the inherent difficulty of determining in what proportion an observed error is due to the dynamic behavior of the process, to model uncertainties, or to measurement uncertainties. Some research in dynamic data reconciliation has been performed by Albuquerque and Biegler (1996), Gertler (1979), Hodouin et al. (1996), Karjala and Himmelblau (1996), Liebman et al. (1992), Makni and Hodouin (1994), McBrayer and Edgar (1994), and Tjoa and Biegler (1991). The dynamic data reconciliation, however, is beyond the scope of the present article, since it is concerned with steady-state data reconciliation.

Standard data reconciliation generally uses numerical optimization procedures that require appreciable computing time for the solution of the nonlinear data reconciliation of a complex process. To render the data reconciliation more suitable for on-line applications, various recursive schemes have been proposed (Hodouin et al., 1996). A recursive algorithm accounting for the trade-off between modeling and measurement errors, similar in this account with the Kalman filtering method, was derived and applied in data reconciliation of mineral processing (Makni and Hodouin, 1994). As an alternative tool to deal with on-line data-processing problems, artificial neural networks have recently found applications in the area of data rectification due mainly to their potential to rapidly and easily perform nonlinear system identification solely on the basis of historical data (Karjala and Himmelblau, 1994). Among them, the use of artificial autoassociative neural networks (AANNs) is particularly of interest in the sense that it is trained with only measurements while simultaneously being able to perform various data-screening tasks, including noise reduction, replacement of missing sensor values (Dong and McAvoy, 1995; Kramer, 1992), and reducing data dimensionality (Kramer, 1991; Tan and Mavrovouniotis, 1995). The autoassociative neutral network is usually a feedforward neural network made of five layers: an input layer, three hidden layers, and one output layer. The function to be mapped is the identity function in the sense that the output layer is identical to the input layer. The first hidden layer is the mapping layer, the second hidden layer having a restricted number of neurons is the bottleneck layer and the third one is the demapping layer. The compression of information via the bottleneck layer results in the aquisition of a correlation model of the input data, thus making it useful to perform the various data-screening tasks. Furthermore, robust autoassociative neural networks (RANNs) were proposed on the basis of the AANN (Kramer, 1992). RANN was shown to have the ability to implement the detect-identify-replace cycle in a single pass for gross error removal while still keeping all the data-processing capacities of the AANN. It is important to note that robustness to gross errors is not automatically a property of autoassociative neural networks. To achieve the objective just cited, the training data set must be augmented by new examples for the gross error inputs and the desired values for the network outputs.

The present study is focused on the development of a novel autoassociative neural network and its application in nonlinear steady-state data reconciliation for a quasi-stationary process, that is, a process for which the process variables exhibit variations of constant variances around constant average values and the process undergoes a very slow dynamics such that the system balances are nearly satisfied at all times.

What makes the novel network different from the traditional autoassociative neural networks lies in the use of a system balance-related least-square error criterion in order to perform simultaneously the systematic error elimination, data smoothing (measurement noise compensation), correction (model deviation compensation), and estimation of the values for unmeasured variables of a stationary process. The system balances of a process may consist of material balances, energy balances, and some others, depending on the nature of the system.

This article is organized as follows. First, the flotation circuit used to generate the data for the present investigation, its modeling, and simulation are briefly introduced. General criteria and methods for data reconciliation are then presented. Finally, the novel neural-net-based method for data rectification of the flotation circuit is presented. A performance comparison study is also made between the novel neural-net approach, AANN, supervised artificial neural networks (SANNs), simple averaging, and standard nonlinear optimization methods. A detailed description and derivation of the system balance-related autoassociative neural network (SBANN) is given in the Appendix.

Brief Description of the Floating Circuit

Flotation is a selective process whereby particles are separated according to their mineral composition. This process is commonly used in mineral-processing industries, and it is usually carried out in flotation cells or flotation columns. To evaluate and compare the performances of the system balance-related autoassociative neural network (SBANN) for nonlinear steady-state data reconciliation, a simulated flotation circuit for the separation of chalcopyrite is used in this study. As shown in Figure 1, the flotation circuit consists of the rougher and scavenger sections, each one consisting of three flotation cells. A complete dynamic simulator of this flotation circuit has been developed. In the derivation of this model, the following fundamental assumptions were made: (1) the pulp in each flotation cell is well mixed; (2) the mechanisms of flotation and entrapment follow first-order kinetics; and (3) each class of particle is made of slow and fast-floating particles as well as nonfloating particles. For a more complete description of the model, the reader is referred to Makni and Hodouin (1994).

This dynamic simulator was used to generate the data that were used to evaluate the steady-state data-reconciliation methods. The process was considered to operate around a stationary point of operation. However, to represent more closely the industrial context, the feed flow rate and composition of the flotation circuit were deliberately and randomly perturbed to create slow dynamic excursions in the vicinity of the stationary point. It was assumed that the mass flow rate of the input and of the two output streams, as well as their mineral composition, were measured with a sampling period of 6 min. In addition, a normally distributed random noise with a standard deviation of 2% of the steady-state value was added to both the flow rate and the mineral composition. A total of 500 sets of measurements corresponding to 50 h of operation were generated. The data were split into training and validation data sets containing, respectively, 400 and 100 data series of points.

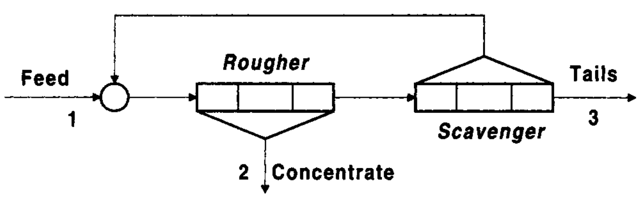

Figure 1. Flotation circuit.

General Data-Reconciliation Criterion

The reconciliation of data, based on observed data and on the model of the process, is normally carried out through the minimization of a least-squares criterion that encompasses both measurement and model uncertainties. The criterion was explained in detail in Hodouin et al. (1996). For the present study, the criterion can be simplified as

$$J(t) = \sum_{m=0}^{M} \left[e_{mt}^T P_m e_{mt} + \sum_{k=t-K}^{t} r_{mkt}^T Q_m r_{mkt} \right] \tag{1}$$

where the terms e are the model uncertainties, that is, the imbalances or the constraint residuals for the main ore phase ($m=0$) and for each of the species ($m=1$ to M), which are the various minerals present in the ore entering the flotation cell. They are evaluated using the following equation:

$$\Gamma \hat{F}_t \hat{x}_{mt} = e_{mt} \tag{2}$$

where Γ is the plant incidence matrix; $\hat{F}_t$ the vector of reconciled ore flow rates at time t; and $\hat{x}_{mt}$ the reconciled composition vector at time t ($x_{0t}=1$ for the ore phase). The r values of Eq. 1 are the measurement residuals defined by the following equations:

$$r_{mkt} = F_k - \hat{F}_t \qquad \text{for } m=0$$

$$r_{mkt} = x_{mk} - \hat{x}_{mt} \qquad \text{for } m=1 \text{ to } M \tag{3}$$

where F_k and x_{mk} are the measured values of flow rates and compositions at time k.

The tuning parameters of the criterion are the weights (P_m and Q_m) and the length of the filtering horizon K. The specifications of the two weight matrices are based on the level of confidence that one has on the material balance model and the measurements, respectively. The length of the filtering window will depend largely on the level of noise and on how stationary the process is. A larger value of K means that the variables are assumed to be operating close to stationary conditions and with a high level of noise. This objective criterion must be solved at each sampling instant using a nonlinear optimization method because of the bilinear or nonlinear nature of the constraints. Depending of the length of the filtering horizon, the numerical optimization technique may be very time-consuming. To render data reconciliation more suitable for on-line applications, various recursive schemes have been proposed (Hodouin et al., 1996). As an alternative, this article examines the use of an autoassociative neural network to perform the function of data reconciliation. This neural network can be viewed as an emulator of the nonlinear reconciliation problem formulated earlier.

Neural-Network Approaches

Supervised training scheme

To develop a neural-network model for steady-state data reconciliation, a supervised training scheme can be adopted (Terry and Himmelblau, 1993). A possible supervised training scheme is depicted in Figure 2. Supervised learning of feedforward neural networks requires that the targets or the desired outputs be specified for each training example. However, in data reconciliation, the true or reconciled outputs of the process are not known and cannot be readily specified. The output target can be generated using the best off-line steady-state reconciliation method along with a representative data set. The neural network, trained with this prereconciled data set, then becomes an emulator of the reconciliation method. The advantage of this procedure is that the neural network can be used more conveniently for on-line application without requiring much calculation. The disadvantage is the need to retrain the neural network if the process conditions change significantly, which may not always be easy to detect.

Since the feedforward neural networks discussed in this article are designed to deal with steady-state data reconciliation problems, the neural network thus performs a static nonlinear mapping of the network input space to the output space. If significant dynamics is present in the process variables to be reconciled, temporal information would have to be considered, otherwise the network would provide the average behavior of the process. In other words, the relevant historical data sets would have to be used in order to predict the values of the reconciled data at each sampling instant.

Autoassociative artificial neural networks

Autoassociative neural networks have been used for steady-state data reconciliation by Kramer (1992). Dong and McAvoy (1995) used an identical network, but the five layers were segmented into two three-layer networks that were trained separately. In these two articles, the transfer function in the mapping and demapping layers was nonlinear, whereas it was linear in the bottleneck and output layers. Figure 3 presents the architecture of the autoassociative neural network used in this investigation, where the inputs (F_1, F_2, F_3, X_1, X_2, and X_3) are the scaled process variables. This neural network simply attempts to replicate the input vector by minimizing the sum of squares of the deviations between the input and the output vectors. Because the number of neurons in the bottleneck layer is significantly smaller than both the input and output layers, a compact representation is therefore obtained. The underlying correlation and redundancies in the input data are thus captured (Kramer, 1992).

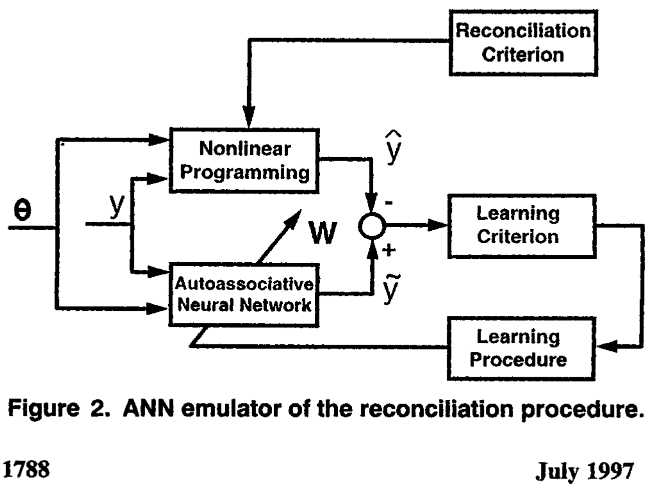

Figure 2. ANN emulator of the reconciliation procedure.

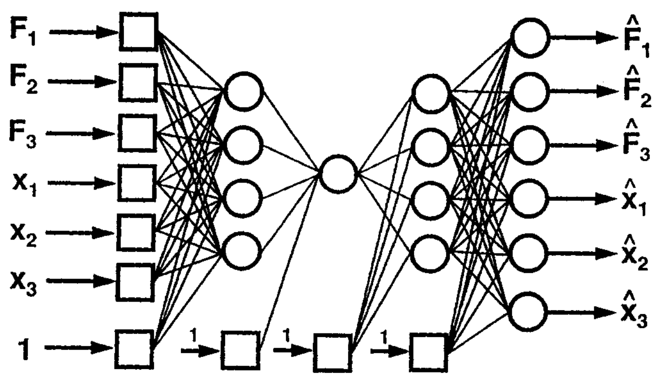

Figure 3. Architecture of the autoassociative neural networks.

In data-reconciliation applications, the bottleneck layer might perform the two expected functions of a reconciliation method: first filter out the measurement noise present in the data by conserving only the essential characteristics necessary to represent adequately the data, and then capture the mass-balance correlation or constraint that exists among the input data. Consequently, it should be possible to select a network architecture containing the required number of neurons in the bottleneck layer that is able to reconcile data, if it is properly trained. If the neural network is trained using a sufficiently large number of representative data patterns, it can be used to produce the adjusted values of the measurements (the outputs of the network) when the raw measurements (the inputs of the network) are available.

To train an autoassociative neural network, a measured data vector is used for both the input and the output of the network, and the objective function to be minimized is nothing more than the sum of squares of the errors between the input and the predicted output vectors. As indicated by Kramer (1992), an AANN is able to remove the uncorrelated measurement noise between the sensors, and to capture the underlying mass-balance correlation. In practice, however, the random errors observed between the sensors may not always be completely uncorrelated. Sharing the same instruments to measure a number of process variables, a situation encountered frequently in mineral processing industries, is one source of correlated measurement errors.

The objective function of an AANN does not contain explicitly the information related to the system balances such as the mass and energy conservation equations of a process. Therefore, it should not be expected that, without using such redundant information, the use of an AANN for data reconciliation will remove or reduce significantly the system imbalance terms in the presence of systematic errors, such as biases, and/or the situation in which the estimation of the values of unmeasured process variables is required.

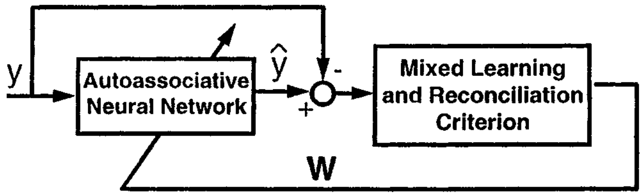

Figure 4. SBANN emulator of the data-reconciliation method.

System balance-related autoassociative neural network

As discussed previously, the supervised training procedure for data reconciliation requires the minimization of two criteria: a reconciliation criterion, and then a learning criterion. It is possible to perform data reconciliation in a single step if the two minimization criteria are combined. This scheme is depicted in Figure 4. The proposed learning criterion of this modified autoassociative neural network contains both the properties of a learning criterion and a reconciliation criterion. The following objective function, like Eq. 1, can be defined as

$$E = \sum_{p=1}^{P} E_p = \frac{1}{2} \sum_{p=1}^{P} \left[\sum_{n=0}^{N} \alpha_n (\hat{y}_{pn} - y_{pn})^2 + \sum_{m=0}^{M} \beta_m \left(g_m \left(\hat{y}_{p1}, \hat{y}_{p2}, \ldots, \hat{y}_{pN} \right) \right)^2 \right] \quad (4)$$

where P is the number of training examples in the training data set; N is the number of process measurements to be reconciled, that is, the length of the composite measurement vector y made of the flow rate and composition vectors; $\hat{y}$ is the output vector of the network, that is, the reconciled composite vector; $g_m(\cdot)$ contains the constraint equations of the mass conservation for flow rates ($m = 0$) and for each species ($m = 1, \ldots, M$); and parameters α_n and β_m are weighting factors. The first term of Eq. 4 therefore represents the weighted sum of squares of the differences between the estimated and measured process variables. The second term represents the weighted sum of squares of the material conservation imbalances, that is, a penalty function for not satisfying the mass-balance equations or other process constraints.

It is obvious that the learning criterion of Eq. 4 cannot exactly reproduce the standard reconciliation method, and a direct comparison can only be made through experimentation. In conventional reconciliation methods, all the temporal information in the window of the past K process measurements is used to filter out the noise and to estimate at each sampling instant the current reconciled values of each process variable. In addition, the mass-conservation imbalances are evaluated strictly with the current process variables. In an SBANN, the two tasks are performed simultaneously with the average knowledge (architecture and weights) that was contained in the minimization of the criterion of Eq. 4 over a learning horizon P. It is important to note that the same horizon P is used for both the residuals $(y_{pn} - \hat{y}_{pn})$ and the imbalances (g_m).

In this article, standard feedforward neural networks are used in conjunction with the objective function defined in Eq. 4 in order to rectify process data. The problem then reduces to finding the parameters (weights) of the neural network that minimize this objective criterion. Details of the backpropagation learning algorithm (Rumelhart et al., 1986) for this special objective function are given in Appendix. This feedforward neural network will be referred to as SBANN.

Results and Discussion

A simple illustrative example

A simple illustrative example is first presented to simply illustrate the influence of the SBANN's weighting coefficients on the data reconciliation performance and the simultaneous correction of a systematic gross error. Consider a data set, comprising 200 values of y_1, y_2, and y_3, generated by the following three equations:

$$y_1 = 0.8 \sin(\theta) + \epsilon_1 \quad (5)$$

$$y_2 = 0.8 \cos(\theta) + \epsilon_2 \quad (6)$$

$$y_3 = y_1 + y_2 + 0.2 + \epsilon_3 \quad (7)$$

The three variables y_1, y_2, and y_3 are calculated using a uniformly distributed value of $\theta \in [0, 2\pi)$, and ϵ_1, ϵ_2, and ϵ_3 are measurement random errors with a Gaussian distribution with a standard deviation of 0.1. Note that the variable y_3 is simply the sum of y_1 and y_2 with a constant systematic bias of 0.2.

The neural network used to reconcile the data sets just described consists of three neurons in both the input and output layers and a single middle layer containing five neurons (not including the neurons for the bias in the input and hidden layers). The specific SBANN learning criterion for this problem can be expressed as

$$E = \sum_{p=1}^{P} E_p = \frac{1}{2} \sum_{p=1}^{P} \left[\sum_{n=1}^{3} \alpha_n (\hat{y}_{pn} - y_{pn})^2 + \beta \left(\hat{y}_{p1} + \hat{y}_{p2} - \hat{y}_{p3} \right)^2 \right] \quad (8)$$

To implement the neural-network-based emulator for data reconciliation, it is important to evaluate the sensitivity of the weighting factors α_1, α_2, α_3, and β through extensive simulations. Figure 5 shows variations in the sum of squares of the

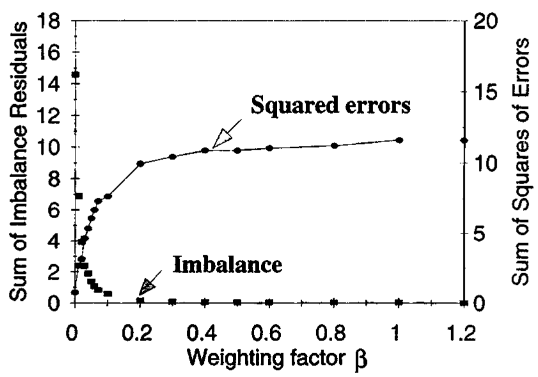

Figure 5. Effect of the weighting factor β on the neural-network performance.

imbalance residuals and the cumulative sum of squares of the errors for the three variables when the parameter β was varied from 0.0 to 1.2 while parameters α_1, α_2 and α_3 were kept constant and equal to, respectively, 0.5, 0.5, and 0.1. For a value for β of zero, the three measurements are assumed to be accurate, so there has been no attempt to correct the system imbalance. As a result, the neutral network has done reasonably well in removing the noise from the original data, but is unaware of the presence of a systematic bias in one of the variables. As the value of β is increased, more weight is placed on the imbalance term and a significant correction of the original data, mainly on y_3, is seen. For a value of β greater than 0.2, the imbalance is almost completely resolved. In general, though, selection of the weighting factors has a certain degree of subjectivity, as the choice depends on the level of confidence one has about the model and the accuracy of the measurement. For the present example, the value of parameter β would be chosen between 0.02 to 0.3.

Figure 6a presents the plots of the noise-free and reconciled data for the following weighting factors: $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $\alpha_3 = 0.1$, and $\beta = 0.2$. These results clearly show that the reconciled values of variables y_1 and y_2 are very close to their original true values, whereas y_3 is adjusted significantly such that the system imbalance term of Eq. 8 becomes quite small, as shown in Figure 6b. It is obvious from this example that the SBANN was able to reduce significantly the systematic bias on variable y_3. This was made possible only because the balance redundancy information was used and significantly less weight was placed on the estimation of variable y_3. If equal weights had been used for the three process variables, the imbalance would have been distributed equally on all three of them. In order to achieve proper data reconciliation, it is therefore important that systematic biases be detected so that the weighting coefficients of the objective criterion can be properly selected or corrected before performing data reconciliation.

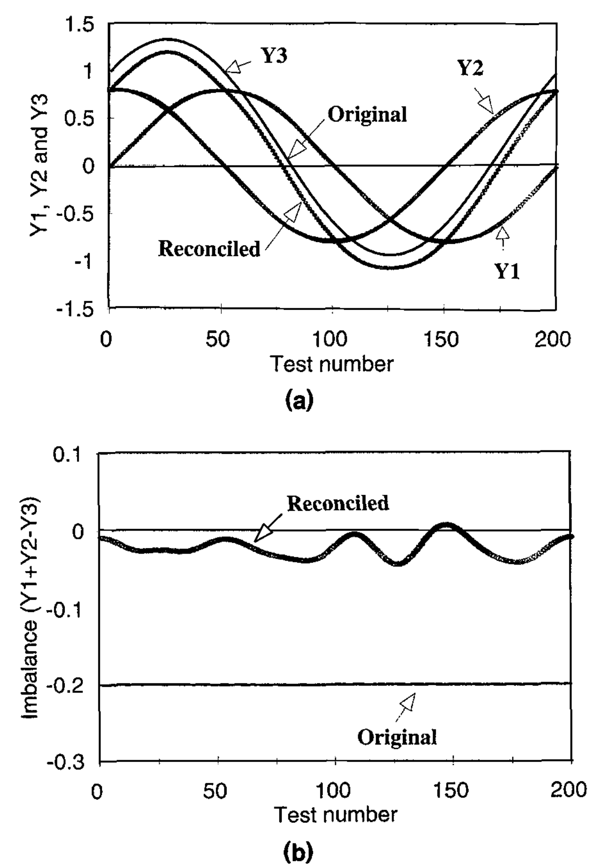

Figure 6. Data reconciled using SBANN for a deterministic process.

(a) Original and reconciled data of Y_1, Y_2 and Y_3; (b) imbalance term relative to the original and reconciled data.

In the next subsection, the performance of a standard five-layer AANN is first examined, and in the subsection following that the ability of a five-layer SBANN to perform proper data reconciliation is evaluated and compared to traditional data-reconciliation methods.

AANN for steady-state data reconciliation

A five-layer AANN similar to the one in Figure 3 is used in this subsection for the reconciliation of the flotation circuit data. The learning data set, consisting of 400 sets of the six measured process variables (the flow rates and compositions in points 1, 2, and 3 of Figure 1) in the absence of nonrandom noise, generated by the flotation circuit simulator, was used to determine the weights of the neural network that minimize the sum of squares of the errors for the identity mapping. An AANN with six inputs and outputs, two neurons in both the mapping and demapping layers and a single neuron in the bottleneck layer, was used. The performance of the AANN was evaluated by using a validation data set composed of 100 additional data measurements. Since, after being properly designed and trained with representative data, the AANN is able to capture the underlying correlation contained in the measurements and to partly exclude the uncorrelated measurement effects, it is thus possible to produce the closest consistent noise-free state for the given input in the least-squares sense. Note that, in general, a simple variance reduction of measurement random noise does not guarantee that the measured data are corrected to satisfy material balances. To better assess these two facets of data reconciliation, the following performance indicators are defined:

$$X_1 = \frac{1}{3P} \sum_{p=1}^{P} \sum_{k=1}^{3} \left(\hat{F}_{kp} - F_{kp}\right)^2 \tag{9}$$

$$X_2 = \frac{1}{3P} \sum_{p=1}^{P} \sum_{k=1}^{3} (\hat{x}_{kp} - x_{kp})^2 \tag{10}$$

$$X_3 = \frac{\sum_{p=1}^{P} \sum_{k=1}^{3} \left(\hat{F}_{kp} - F^*_{kp}\right)^2}{\sum_{p=1}^{P} \sum_{k=1}^{3} (F_{kp} - F^*_{kp})^2} \tag{11}$$

$$X_4 = \frac{\sum_{p=1}^{P} \sum_{k=1}^{3} (\hat{x}_{kp} - x^*_{kp})^2}{\sum_{p=1}^{P} \sum_{k=1}^{3} (x_{kp} - x^*_{kp})^2} \tag{12}$$

$$X_5 = \frac{1}{P} \sum_{p=1}^{P} \left(\hat{F}_{1p} - \hat{F}_{2p} - \hat{F}_{3p} \right)^2 \quad (13)$$

$$X_6 = \frac{1}{P} \sum_{p=1}^{P} \left(\hat{F}_{1p}\hat{x}_{1p} - \hat{F}_{2p}\hat{x}_{2p} - \hat{F}_{3p}\hat{x}_{3p} \right)^2 \quad (14)$$

Criteria X_1 and X_2 account for the deviations between the corrected flow rates and compositions and their experimental values. These are the only two criteria that the AANN attempts to minimize. With simulated data, the various methods can also be evaluated for their ability to estimate the true values of each variable. This performance evaluation is given by criteria X_3 and X_4. The closer these indicators are to zero, the closer the estimated values are to their true values in the sense of the least-square error. A value of one indicates that the measured data were not corrected. Of course, for industrial data, it is not possible to calculate these criteria, because the true values of F_{kp}^* and x_{kp}^* are never available. Finally, criteria X_5 and X_6 provide an evaluation of the ability of the data-reconciliation methods to correct flow rate and composition imbalances.

Table 1 gives the six performance indicators for the autoassociative neural network (AANN) for both the training (T) and the validation (V) data sets. These results clearly show that the AANN was able to satisfactorily reconcile all the process variables as given by the criteria X_3 and X_4. This was made possible because the AANN was able to remove a significant portion of the measurement errors, that is, the random Gaussian noise. As a result, the material imbalances (X_5 and X_6) were also significantly reduced because, in this case, the imbalances mainly come from the presence of the white noise on each of the process variables. Therefore, the use of AANN is justified for data reconciliation in this particular situation.

In the next section, a five-layer SBANN is trained and tested using (1) the same data sets used for the AANN, that is, data containing only zero-mean Gaussian noise; (2) the identical data sets used in the previous example, except for the addition of a constant bias to the measurement of the tail mass flow rate; and (3) the identical data sets used with the AANN, except for the presence of one unmeasured process variable that will be estimated simultaneously when all the measured variables are reconciled.

SBANN for steady-state data reconciliation

The structure of the SBANN, used in this section, for steady-state data reconciliation is the same as the one used for the AANN. It is simply a five-layer feedforward neural network that attempts to perform the identity mapping between the inputs and outputs of the neural network. The mapping and demapping layers contain two neurons and the bottleneck layer contains a single neuron.

The specific objective function to be minimized for SBANN is defined by the following equation:

$$E = \frac{1}{2} \sum_{p=1}^{P} \left[\left(\sum_{k=1}^{6} \alpha_k (\hat{y}_{pk} - y_{pk})^2 \right) + \beta_0 \left(\hat{y}_{p1} - \hat{y}_{p2} - \hat{y}_{p3} \right)^2 + \beta_1 \left(\hat{y}_{p1}\hat{y}_{p4} - \hat{y}_{p2}\hat{y}_{p5} - \hat{y}_{p3}\hat{y}_{p6} \right)^2 \right] \quad (15)$$

where pairs ($\hat{y}_{p1}$, $\hat{y}_{p4}$), ($\hat{y}_{p2}$, $\hat{y}_{p5}$), and ($\hat{y}_{p3}$, $\hat{y}_{p6}$) stand for the adjusted flow rate and mineral composition in the feed, the concentrate, and the tails, respectively. Parameters α_k, β_0, and β_1 are weighting factors. The objective criterion, defined in Eq. 15, is significantly different from the objective criterion of Eq. 1, even though the final objective is the same. The main difference lies in the way the filtering of the data is being performed. In Eq. 1, filtering is achieved by the summation over a moving window of length K of the past measurements, whereas in Eq. 15 the filtering is performed by the neural network itself and by the learning horizon P. When the backpropagation algorithm is used, caution must be exercised in specifying the weighting factors, α_k, β_0, and β_1 in Eq. 15, because of the magnitude of the output errors that are backpropagated through the neural network. To avoid possible convergence problems and to maintain an efficient learning rate, these parameters should be chosen based on the knowledge of the measurement error level, the model error, and the setting of the learning-rate parameter, η. When the quasi-Newton optimization procedure is used, this restriction does not apply.

In general, the steady-state mass-balance equations of such quasi-stationary processes were approximately satisfied at any given instant, therefore the weighting parameters β_0 and β_1 were simply given sufficiently large values. Obviously, when

Table 1. Performance Comparison of Steady-State Data Reconciliation

	X_1 t^2h^{-2}		X_2 Mass Fraction		X_3 (—)		X_4 (—)		X_5 t^2h^{-2}		X_6 t^2h^{-2}	
Method	T	V	T	V	T	V	T	V	T	V	T	V
Original data	0	0	0	0	1	1	1	1	8.7	6.3	1.4×10^{-2}	1.5×10^{-2}
AANN (6-2-1-2-6)	2.39	1.99	5.1×10^{-6}	4.2×10^{-6}	0.23	0.22	0.13	0.12	6.3×10^{-1}	1.4×10^{-1}	8.5×10^{-4}	1.8×10^{-4}
SBANN (6-2-1-2-6)												
$\alpha_{1-6}=1.0$, $\beta_{1,2}=0.5$	2.48	2.24	5.4×10^{-6}	4.4×10^{-6}	0.17	0.16	0.10	0.10	4.8×10^{-4}	3.5×10^{-4}	4.3×10^{-5}	3.0×10^{-5}
Nonlinear optim ($K=1$)	2.26	2.02	4.5×10^{-6}	3.3×10^{-6}	0.22	0.25	0.33	0.36	1.6×10^{-7}	1.1×10^{-7}	7.8×10^{-7}	9.6×10^{-7}
Nonlinear optim ($K=4$)	2.60	2.40	4.9×10^{-6}	4.7×10^{-6}	0.11	0.10	0.17	0.16	3.2×10^{-7}	2.5×10^{-7}	1.7×10^{-6}	2.5×10^{-6}
Nonlinear optim ($K=8$)	2.73	2.38	5.4×10^{-6}	4.8×10^{-6}	0.06	0.06	0.10	0.10	7.5×10^{-7}	6.2×10^{-7}	3.1×10^{-6}	7.8×10^{-6}
Average (total)	3.88	2.96	7.0×10^{-6}	5.4×10^{-6}	0.38	0.14	0.02	0.09	1.2×10^{-4}	4.4×10^{-2}	3.9×10^{-6}	3.2×10^{-4}
SANN-1	2.56	2.20	5.9×10^{-6}	4.7×10^{-6}	0.09	0.10	0.05	0.07	1.7×10^{-3}	2.4×10^{-3}	1.4×10^{-6}	2.2×10^{-6}
SANN-2	2.48	2.13	5.2×10^{-6}	4.6×10^{-6}	0.11	0.12	0.09	0.10	2.2×10^{-3}	2.5×10^{-3}	2.7×10^{-6}	3.1×10^{-6}

β_0 and β_1 in Eq. 15 are both set to zero, the SBANN becomes an ordinary AANN, which aims only to minimize the sum of the squares of the errors between the measurements and the adjusted data without explicitly taking the information related to the mass conservation redundancy into account.

Table 1 presents the results obtained with the five-layer SBANN trained with the identical data sets that were used in the previous section. When compared to the results obtained with the AANN, the SBANN performs slightly better. The inclusion of the imbalance terms in the objective cost function to be minimized makes it possible for the SBANN to use this redundancy information directly to filter out the portion of the correlated random errors between the measurements, and to reduce the uncorrelated measurement errors. On the other hand, AANN is only able to remove the uncorrelated random noise between the sensors. Note that the location of the correlated random errors between the sensors can be encountered frequently in mineral processing industries, for example, in the case where the same instruments are used to measure a number of process variables. In the simulation, the introduction of random noise in the feed flow rate is mainly responsible for a portion of correlated random errors.

To better assess the ability of the AANN and the SBANN to perform steady-state data reconciliation efficiently relative to a conventional data-reconciliation method, an unconstrained optimization was performed to minimize the objective function of Eq. 1 using the same data sets used for the two neural networks. The weighting matrix, Q_m, in Eq. 1 accounts for the relative confidence that one has in the measurements of each of the variables. If no systematic biases are detected, this level of confidence is usually associated with the covariance matrix, and Q_m is normally equal to the inverse of the process variable covariance matrix. P_m is set to a value that corresponds to the relative confidence level on the models, expressed here as material balances. Like the specification of the weight vectors in the SBANN, P_m was set to a diagonal matrix whose elements were relatively large because of the quasi-stationary nature of the process and therefore a relatively high confidence level in the mass-balance model. The results obtained for data windows of lengths 1, 4, and 8 are presented in Table 1. As expected, the reconciled data using the optimization method are closer to their true values when the length of the data window is increased ($K = 8$). The nonlinear optimization method for this particular data set performs best in terms of estimation accuracy when the length of the data window is sufficiently long to efficiently filter out the measurement noise. For this particular data-reconciliation method, there is of course no distinction between the training and the validation data sets, since the optimization procedure must be performed independently at each sampling instant. This method has the disadvantage of requiring significantly more computational effort but, on the other hand, has the advantage of using local information and, as a result, it is more adapted to processes that are not entirely stationary, as in the present situation. When properly trained, it can be readily used, without further training, for the SBANN, to estimate each of the process variables, but the information contained in the weights and the structure of the network represents the average trend that is expressed by the training data set.

Since steady-state conditions were assumed, it is also interesting to evaluate the reconciliation performance using a simple average of all measured variables over the entire length of the data sets, that is, the 400 and 100 values, respectively, for the training and validation data sets of each mass flow rate and composition. If the process were perfectly under constant steady-state conditions and only Gaussian noise was present, the simple averaging of all process variables, over a sufficiently large data window, should lead to an excellent data reconciliation. The results of the pure data filtering are also presented in Table 1. It was observed, however, that the criterion X_3 for the training data samples is considerably higher than those obtained with most of the other data-reconciliation methods. Pure filtering, or the simple averaging method, was not successful for the steady-state data reconciliation of this particular process in which a slow dynamics exists and the true process variables of the three mass flow rates are not constant but vary quite significantly with time.

Results obtained with a SANN are also presented in Table 1. SANN-1 and SANN-2 represent the results obtained with a three-layer feedforward neural network for, respectively, training data sets of a hypothetically perfect data reconciliation, and one that corresponds to the output of the nonlinear optimization procedure with a horizon K of 8. These results clearly show that an ordinary three-layer feedforward neural network could be used to emulate with relatively high accuracy the performance of the best data reconciliation, and then could be used on-line more readily than any other data-reconciliation method. The fact that other results (not shown here), obtained with SANN-1 for a truly stationary process, led to nearly perfect behavior suggests that the data-reconciliation results of Table 1 for the SANN-1 are the best that could be expected from neural networks for the particular quasi-stationary process used in this study. Indeed, without the presence of temporal information, the neural network can only represent the average behavior expressed by the data set.

Each data-reconciliation method will now be evaluated for their ability to perform steady-state data reconciliation in the case where a constant bias is present on one of the measured process variables. Table 2 shows the performance of each data-reconciliation method for the data sets that are identical to those used earlier, except that the tail mass flow rate measurement is affected by a constant bias. Of course, this bias must be detected before performing data reconciliation in order to adjust the weighting coefficients; otherwise, this constant bias will be distributed to all the process variables in proportion to the inverse of their weighting coefficient. In this evaluation, the bias was assumed to be approximately 5% of the average value of the tail flow rate. Accordingly, the weighting factor, α_3, associated with the tail mass flow rate term was given a value of 0.1 instead of the weighting coefficient of 1.0 for other measured variables. The results in Table 2 clearly show that the SBANN performs quite well. The performance of both the AANN and the total average filtering method, which cannot account for the presence of a systematic bias in the data set, is totally inadequate. The results in Table 2 show that the nonlinear optimization method is still the best one in terms of estimation accuracy.

Finally, the ability of the SBANN and the nonlinear optimization method are now evaluated in the case where one of

Table 2. Performance Comparison of Steady-State Data Reconciliation in the Presence of Bias

Method	X_1 t^2h^{-2}		X_2 Mass Fraction		X_3 (—)		X_4 (—)		X_5 t^2h^{-2}		X_6 t^2h^{-2}	
	T	V	T	V	T	V	T	V	T	V	T	V
Original data	0	0	0	0	1	1	1	1	3.5×10^{1}	3.3×10^{1}	1.4×10^{-2}	1.2×10^{-2}
AANN (6-2-1-2-6)	1.59	1.67	4.9×10^{-6}	6.2×10^{-6}	0.87	0.83	0.15	0.18	2.9×10^{1}	2.3×10^{1}	3.9×10^{-3}	1.6×10^{-3}
SBANN (6-2-1-2-6)												
$\alpha_{1,2,4-6}=1.0$, $\alpha_3=0.1$ $\beta_{1,2}=0.5$	8.96	8.39	4.2×10^{-6}	5.6×10^{-6}	0.06	0.06	0.15	0.20	1.2×10^{-2}	1.2×10^{-2}	6.3×10^{-4}	5.7×10^{-4}
Nonlinear optim ($K=4$)	11.14	10.05	4.2×10^{-6}	4.4×10^{-6}	0.05	0.05	0.18	0.17	2.3×10^{-9}	3.6×10^{-5}	1.1×10^{-4}	1.2×10^{-4}
Nonlinear optim ($K=8$)	11.09	10.12	4.8×10^{-6}	4.3×10^{-6}	0.03	0.03	0.10	0.16	9.2×10^{-11}	9.1×10^{-11}	2.6×10^{-6}	2.1×10^{-6}
Average (total)	3.85	3.46	6.1×10^{-6}	5.7×10^{-6}	0.84	0.77	0.21	0.06	2.5×10^{1}	$2.5\times10^{+1}$	7.8×10^{-6}	9.1×10^{-5}

the process variables is not measured and needs to be estimated. It is important to note that the AANN cannot be simply used to estimate the values of unmeasured variables. To test these two methods, it was assumed that the tail mass flow rate, y_{p3}, of the flotation circuit could not be measured. Thus, to account for the unmeasured process variable, two modifications of the SBANN are required: first, the weighting coefficient, α_3, in the objective cost function (Eq. 15) has to be set to zero; and second, the structure of the neural network has to be modified. Since there are now only five process measurements available, the SBANN input layer contains only five neurons, excluding the bias neuron. The output layer of the neural network still contains six neurons for the six adjusted process variables, among which one is not measured. The network structure is thus designed to be (5-2-1-2-6) and it is trained with the identical data set used with the AANN, except for the absence of the tail mass flow rate, y_{p3}. In the neural network training procedure, the initial values of $\hat{y}_{p3}$ can be arbitrarily set. However, to promote faster convergence, it is better to set its value to the estimated average value of this variable, as was done in this study. The performance indices for the two reconciliation methods are presented in Table 3. Figure 7a presents the plots of the simulated and true values of the unmeasured tail mass flow rate. Note that the simulated unmeasured tail mass flow rate is only used for comparison with the SBANN performance of the estimation, and does not exist in real situations. The true values of the unmeasured variable are the simulated values without the measurement random noise included. This graph allows visualization of the quasi-stationary behavior of the process and the level of noise on process variables. Figure 7b presents the plots of the true tail flow rate and the estimated values obtained with the SBANN. These results show that the SBANN is able to perform proper data reconciliation and, at the same time, to estimate an unmeasured process variable. As expected, the nonlinear optimization method also performs extremely well for this example. It is important to point out that the estimation of an unmeasured variable is possible only if data redundancy is present or, in other words, if the unmeasured process variable is observable. It is also noticed that, even though the off-line training of the SBANN used in this study can take a few minutes to converge on a UNIX IBM 320 workstation, the on-line use of the neural networks for data reconciliation only requires a very short period of computing time because a very simple algebraic calculation is involved. The computing time for data reconciliation using nonlinear optimization methods may be appreciable, depending on the complexity of the problems. In this study, an unconstrained optimization algorithm using a simplex search method was used to perform the nonlinear steady-state data reconciliation. The maximum number of iterations was set to 1,200. The average computing time using the optimization method for each set was about 3 s.

Conclusion

A novel autoassociative neural-network-based estimator (SBANN) was developed to cope with nonlinear steady-state data-rectification problems. The objective function of the SBANN is the same as that used in traditional autoassociative neural networks (AANN), except that a system mass-balance penalty function term is included, thus making the AANN a special case of the SBANN. The SBANN conserves all the data-reconciliation capabilities of the AANN. In addition, because of the explicit use of the redundancy information related to system mass balances, use of the SBANN is also able to adjust the measured data to make them as consistent as possible with the system mass balance model, especially in the presence of nonrandom measurement errors. It

Table 3. Performance Comparison of Steady-State Data Reconciliation for an Unmeasured Variable

Method	X_1 t^2h^{-2}		X_2 Mass Fraction		X_3 (—)		X_4 (—)		X_5 t^2h^{-2}		X_6 t^2h^{-2}	
	T	V	T	V	T	V	T	V	T	V	T	V
Original data	0	0	0	0	1	1	1	1	5.6	4.5	1.4×10^{-2}	1.5×10^{-2}
SBANN (5-2-1-2-6)												
$\alpha_{1,2,4-6}=1.0$, $\alpha_3=0.0$ $\beta_{1,2}=0.5$	2.83	2.71	5.5×10^{-6}	4.5×10^{-6}	0.20	0.20	0.14	0.10	1.4×10^{-2}	1.3×10^{-2}	7.6×10^{-4}	6.8×10^{-4}
Nonlinear optim ($K=4$)	2.78	2.60	4.8×10^{-6}	4.4×10^{-6}	0.17	0.16	0.17	0.16	1.2×10^{-9}	1.5×10^{-9}	9.7×10^{-5}	1.2×10^{-4}
Nonlinear optim ($K=8$)	2.84	2.51	5.3×10^{-6}	4.6×10^{-6}	0.09	0.10	0.10	0.09	3.8×10^{-11}	6.0×10^{-11}	2.5×10^{-6}	4.6×10^{-6}

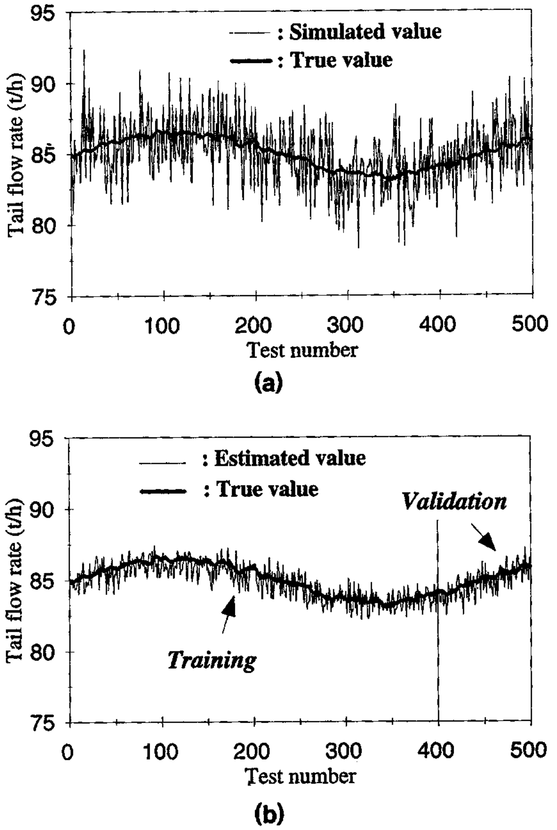

Figure 7. Use of the SBANN to estimate the value of an unmeasured variable.

(a) Comparison of the simulated measurement and the true value; (b) comparison of the estimated and the true values.

is also able to simultaneously estimate the values of unmeasured variables. Simulation studies have been conducted on the application of this novel neural network for data reconciliation of a flotation circuit with good success.

The SBANN-based estimator, developed in this study, exhibited satisfactory filtering capability and robustness. However, it may be necessary to adapt this neural network on-line in order to cope with significant steady-state changes in the system that may occur in practice. One possible method would be to conduct an off-line training of the SBANN in a moving window of past measurements in order to adapt the neural model to changing process conditions.

Work is now in progress for the development of an SBANN-based estimator to be used for dynamic data reconciliation.

Acknowledgment

The project is part of the Knowledge Based Automatic Control (KBAC) sponsored by 14 Canadian mining companies and two governmental research centers. The authors wish to thank these organizations for their support, as well as the Natural Science and Engineering Research Council of Canada.

Notation

e_{mt} = vector of mass residual imbalances, Eq. 2
E = objective cost function, Eq. 4
F = flow rate scaled from 0 to 1, Figure 3
$J(t)$ = objective cost criterion, Eq. 1
M = total number of mineral species
S = output of the hidden neuron
X = mass fraction scaled from 0 to 1, Figure 3
X_i = performance indicator
Z = output of the final layer of neurons
δ = vector of propagated errors through the network

Subscripts and superscripts

m = mth species
n = nth measured variable
p = training sample number
i = weighting factor index number (superscript)
$*$ = true value

Literature Cited

Albuquerque, J. S., and L. T. Biegler, "Data Reconciliation and Gross-Error Detection for Dynamic Systems," *AIChE J.*, **42**(10), 2841 (1996).

Aynsley, M., A. Hofland, A. J. Morris, G. A. Montague and C. Di Massimo, "Artificial Intelligence and the Supervision of Bioprocesses (Real-Time Knowledge-Based Systems and Neural Networks)," *Adv. Biochem. Eng. Biotechnol.*, **48**, 1 (1993).

Chen, J., A. Bandoni, and J. A. Romagnoli, "Robust Statistical Process Monitoring," *Comput. Chem. Eng.*, **20**, S497 (1996).

Crowe, C. M., Y. A. Garcia Campos, and A. Hrymak, "Reconciliation of Process Flow Rates by Matrix Projection: I. The Linear Case," *AIChE J.*, **29**(6), 881 (1983).

Crowe, C. M., "Reconciliation of Process Flow Rates by Matrix Projection: II. The Nonlinear Case," *AIChE J.*, **32**(4), 616 (1986).

Crowe, C. M., "Observability and Redundancy of Process Data for Steady-State Reconciliation," *Chem. Eng. Sci.*, **44**(12), 2909 (1989).

Crowe, C. M., "Data Reconciliation—Progress and Challenges," *J. Proc. Control*, **6**(2/3), 89 (1996).

Dong, D., and T. J. McAvoy, "Nonlinear Principal Component Analysis-based on Principal Curves and Neural Networks," *Comput. Chem. Eng.*, **20**(1), 65 (1995).

Dunia, R., S. J. Qin, T. F. Edgar, and T. J. McAvoy, "Use of Principal Component Analysis for Sensor Fault Identification," *Comput. Chem. Eng.*, **20**, S539 (1996).

Gertler, J., "A Constrained Minimum Variance Input-Output Estimator for Linear Dynamic Systems," *Automatica*, **15**(3), 353 (1979).

Hodouin, D., C. Bazin, and S. Makni, "Dynamic Material Balance Algorithm: Application to Industrial Flotation Circuits," *Proc. SME/AIME Meeting*, Phoenix, AZ, p. 96 (1996).

Hodouin, D., C. Bazin, and S. Makni, "On-Line Reconciliation of Mineral Processing Data," *Proc. AIME/SME Symp. on Emerging Computer Techniques for the Mineral Industry*, Reno, NV (1993).

Hodouin, D., T. Gelpe, and M. D. Everell, "Sensitivity Analysis of Material Balance Calculations: An Application to a Cement Clinker Grinding Process," *Powder Technol.*, **32**, 139 (1982).

Hodouin, D., and M. D. Everell, "A Hierarchical Procedure for Adjustment and Material Balancing of Mineral Process Data," *Int. J. Miner. Process.*, **7**(2), 91 (1980).

Hoskins, J. C., K. M. Kaliyur, and D. M. Himmelblau, "Fault Diagnosis in Complex Chemical Plants Using Artificial Neural Networks," *AIChE J.*, **37**(1), 137 (1991).

Karjala, T. W., and D. M. Himmelblau, "Dynamic Data Reconciliation by Recurrent Neural Networks vs. Traditional Methods," *AIChE J.*, **40**(11), 1865 (1994).

Karjala, T. W., and D. M. Himmelblau, "Dynamic Rectification of Data via Recurrent Neural Nets and Extended Kalman Filter," *AIChE J.*, **42**(8), 2225 (1996).

Kramer, M. A., "Nonlinear Principal Component Analysis Using Autoassociative Neural Networks," *AIChE J.*, **37**(2), 233 (1991).

Kramer, M. A., "Autoassociative Neural Networks," *Comput. Chem. Eng.*, **16**(4), 313 (1992).

Liebman, M. J., T. F. Edgar, and L. S. Lasdon, "Efficient Data Reconciliation and Estimation for Dynamic Processes Using Nonlinear Programming Techniques," *Comput. Chem. Eng.*, **16**(10/11), 963 (1992).

Lynch, A. J., *Mineral Crushing and Grinding Circuits: Their Simulation, Optimization and Control, Elsevier, New York, p. 137 (1977).*

MacGregor, J. F., "Statistical Process Control of Multivariable Processes," *Control Eng. Prac.*, **3**(3), 403 (1995).
Mah, R. S. H., *Chemical Process Structures and Information Flows*, Batterworths, Boston (1990).
Makni, S., D. Hodouin, and C. Bazin, "On-Line Reconciliation by Minimization of a Weighted Sum of Squared Residuals and Node Imbalances," *Proc. IMPC*, Publ. SME/AIME, San Francisco (1995).
Makni, S., and D. Hodouin, "Recursive BILMAT Algorithm: An On-Line Extension of Data Reconciliation Techniques for Steady-State Bilinear Material Balance," *Min. Eng.*, **7**(9), 1179 (1994).
McBrayer, K. F., and T. F. Edgar, "Detection and Estimation of Measurement Bias in Dynamic Data Reconciliation," *Advanced Control of Chemical Processes (ADCHEM'94)*, IFAC, Pergamon Press, London (1994).
Meyer, M., B. Koehret, and M. Enjalbert, "Data Reconciliation on Multicomponent Network Process," *Comput. Chem. Eng.*, **17**(8), 807 (1993).
Nomikos, P., and J. F. MacGregor, "Monitoring Batch Processes Using Multiway Principal Component Analysis," *AIChE J.*, **40**(8), 1361 (1994).
Piovoso, M. J., and K. A. Kosanovich, "Application of Multivariate Statistical Methods to Process Monitoring and Controller Design," *Int. J. Control*, **59**(3), 743 (1994).
Rollins, D. K., and J. F. Davis, "Unbiased Estimation of Gross Errors in Process Measurements," *AIChE J.*, **38**(4), 563 (1992).
Romagnoli, J. A., and G. Stephanopoulos, "Rectification of Process Measurement Data in the Presence of Gross Errors," *Chem. Eng. Sci.*, **36**(11), 1849 (1981).
Rumelhart, D., G. Hinton, and R. Williams, "Learning Internal Representations by Error Propagation," *Parallel Distributed Processing*, MIN Press, Cambridge, MA (1986).
Schraa, O. J., and C. M. Crowe, "The Numerical Solution of Bilinear Data Reconciliation Problems Using Unconstrained Optimization Methods," *Comput. Chem. Eng.*, **20**, S727 (1996).
Smith, H. W., and N. Ichiyen, "Computer Adjustment of Metallurgical Balances," *Can. Inst. Mining Met. Bull.*, **66**, 97 (1973).
Tan, S., and M. L. Mavrovouniotis, "Reducing Data Dimensionality through Optimizing Neural Network Inputs," *AIChE J.*, **41**(6), 1471 (1995).
Terry, P. A., and D. M. Himmelblau, "Data Rectification and Gross Error Detection in a Steady-State Process via Artificial Neural Networks," *Ind. Eng. Chem. Res.*, **32**, 3020 (1993).
Tjoa, I. B., and L. T. Biegler, "Simultaneous Strategies for Data Reconciliation and Gross Error Detection of Nonlinear Systems," *Comput. Chem. Eng.*, **15**(10), 679 (1991).
Tong, H., and C. M. Crowe, "Detecting Persistent Gross Errors by Sequential Analysis of Principal Components," *Comput. Chem. Eng.*, **20**, S733 (1996).
Wiegel, R. L., "Advances in Mineral Processing Material Balance," *Can. Metall. Q.*, **11**(2), 413 (1972).
Wise, B. M., and N. L. Ricker, "Recent Advances in Multivariable Statistical Process Control: Improving Robustness and Sensitivity," *Proc. IFAC, ADCHEM Symp.*, 125 (1991).
Zasadzinski, M., "Contribution à l'estimation de l'état des systèmes singuliers. Application à la validation des données des systèmes dynamiques linéaires," PhD Thesis, Univ. de Nancy, France (1990).

Appendix

The backpropagation learning algorithm (Rumelhart et al., 1986) was used in this study for the special objective function, which contains not only the term of the least-square errors but also the process mass imbalances. To derive the learning algorithm, let's look at a standard three-layer feedforward neural network as shown in Figure A1. The output of each neuron of this neural network is calculated as follows. First, the weighted summation of each neuron in the hidden and output layers are evaluated for each pattern, p:

$$s_{pj} = \sum_{i=1}^{I-1} W_{ij} y_{pi} + W_{Ij} \tag{A1}$$

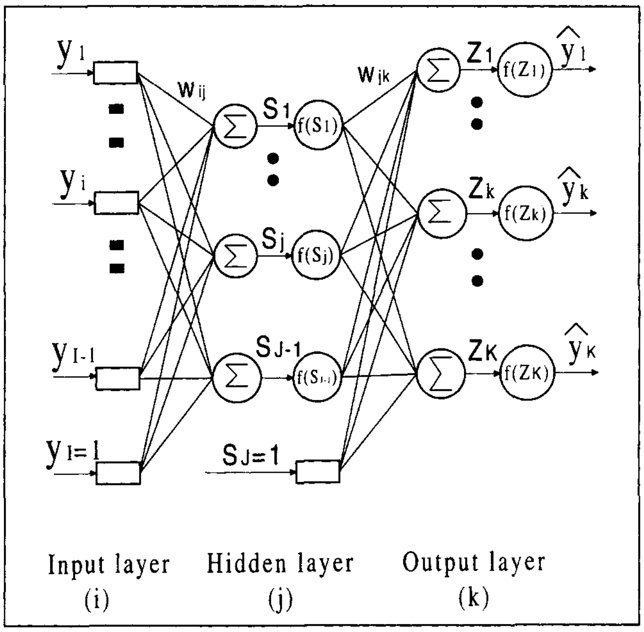

Figure A1. Feedforward neural-network structure.

$$Z_{pk} = \sum_{j=1}^{J-1} W_{jk} f(S_{pj}) + W_{Jk} \tag{A2}$$

where W_{ij} and W_{jk} are the connection weights, respectively, of neuron i to neuron j *and* of neuron j to neuron k.

Next, the values of S_{pj} and Z_{pk} are processed through a transfer function. In this investigation, a sigmoidal transfer function is used:

$$f(S_{pj}) = \frac{1}{1+e^{-S_{pj}}} \tag{A3}$$

$$f(Z_{pk}) = \frac{1}{1+e^{-Z_{pk}}} = \hat{y}_{pk} \tag{A4}$$

The learning rule for multilayer feedforward neural networks is generally the well-known backpropagation algorithm (Rumelhart et al., 1986) that is obtained by differentiating a least-squares objective function with respect to each weight in the neural network:

$$W_{jk}^{(t+1)} = W_{jk}^{(t)} + \eta\left(\frac{\partial E}{\partial W_{jk}}\right) = W_{jk}^{(t)} + \eta\left(\frac{\partial E}{\partial Z_k}\right)\left(\frac{\partial Z_k}{\partial W_{jk}}\right) = W_{jk}^{(t)} + \eta \delta_{pk} f(S_{pj}) \tag{A5}$$

$$W_{ij}^{(t+1)} = W_{ij}^{(t)} + \eta\left(\frac{\partial E}{\partial W_{ij}}\right) = W_{ij}^{(t)} + \eta\left(\frac{\partial E}{\partial S_j}\right)\left(\frac{\partial S_j}{\partial W_{ij}}\right) = W_{ij}^{(t)} + \eta \delta_{pj} f(y_{pi}) \tag{A6}$$

where $W_{ij}^{(t+1)}$ and $W_{jk}^{(t+1)}$ are the weights, evaluated at time t, that will prevail from iteration t to $(t+1)$. The parameter η in Eqs. 20 and 21 is the learning rate. Note that even though these equations were defined for a three-layer neural network, similar equations are easily derived for a five-layer neural network.

The backpropagation algorithm is usually only concerned with the sum of squares of the deviations between the predicted and target values. For the SBANN, the system imbalance term is included in the criterion, along with the sum of squares of the deviations. Therefore, for a three-layer neural network, the output error, δ_{pk}, and the propagated error for the neurons of the hidden layer, δ_{pj}, can then be calculated:

$$\delta_{pk} = \frac{\partial E}{\partial Z_k} = -\frac{\partial E}{\partial \hat{y}_{pk}} \frac{\partial \hat{y}_{pk}}{\partial Z_k}$$

$$= \left[\alpha_k (y_{pk} - \hat{y}_{pk}) + \sum_m \beta_m g_m(\hat{y}_{p1}, \ldots, \hat{y}_{pN}) \frac{\partial g_m(\hat{y}_{p1}, \ldots, \hat{y}_{pN})}{\partial \hat{y}_{pk}} \right] \times \hat{y}_{pk}(1 - \hat{y}_{pk}) \quad \text{(A7)}$$

$$\delta_{pj} = -\frac{\partial E}{\partial S_j} = -\frac{\partial E}{\partial f(S_{pj})} \frac{\partial f(S_{pj})}{\partial S_j}$$

$$= \left(\sum_{k=1}^{K} \delta_{pj} W_{kj} \right) f(S_{pj})[(1 - f(S_{pj}))] \quad \text{(A8)}$$

A backpropagation algorithm for a five-layer SBANN can be determined in the same way as described before. This derivation allows the backpropagation algorithm to be used directly, but also the analytical derivatives to be calculated in the case where the quasi-Newton optimization routine is used.

Manuscript received June 25, 1996, and revision received Mar. 20, 1997.