Instrumentation Network Design and Upgrade for Process Monitoring and Fault Detection

Miguel Bagajewicz and Adrian Fuxman
School of Chemical Engineering, University of Oklahoma, Norman, OK 73019

Ariel Uribe
Universidad Industrial de Santander, Bucaramanga, Colombia

DOI 10.1002/aic.10279
Published online in Wiley InterScience (www.interscience.wiley.com).

A mixed-integer linear programming formulation for the design of sensor networks for simultaneous process monitoring and fault detection/resolution is presented. The objective is to find a cost-optimal sensor set for a chemical process that provides a good estimate of the state of the system and detects as well as locate a preestablished set of faults. A continuously stirred tank reactor example is provided and extensions are discussed. © 2004 American Institute of Chemical Engineers *AIChE J,* 50: 1870–1880, 2004
Keywords: sensors, sensor systems, fault detection, instrumentation upgrade

Introduction

The design and upgrade of instrumentation networks consist in determining the optimal set of instruments such that several goals are fulfilled. Among others, the most important ones are that sufficiently accurate estimates of variables/parameters of interest are obtained and faults are detected and eventually identified. The former are needed for normal operations monitoring and real-time optimization, whereas the latter are essential to guarantee the continuity and safety of operations. The usual objective to minimize is cost.

There are several articles devoted to the study of the design of sensor networks using goals corresponding to normal monitoring operations. Aside from cost, different other objective functions such as precision (Luong et al., 1994; Madron and Veverka, 1992), reliability (Ali and Narasimhan, 1993, 1996), or simply observability (Bagajewicz and Sánchez, 1999a; Luong et al., 1994; Madron, 1992; Maquin et al, 1994) were used. Different techniques were also used, such as graph theory (Ali and Narasimhan, 1993, 1996; Madron, 1992), mathematical programming (Bagajewicz, 1997; Bagajewicz and Cabrera, 2001a), genetic algorithms (Sen et al., 1998; Viswanath and Narasimhan, 2001), multiobjective optimization (Bagajewicz and Cabrera, 2001b; Viswanath and Narasimhan, 2001), among others. Bagajewicz and Sánchez (1999a) showed in detail that there is a connection between the maximum precision and the minimum cost models. This connection states that the solution of one is one solution of the other. Bagajewicz and Sánchez (2000a) also showed that this same type of duality holds for the objectives of reliability and cost. In essence, it is sufficient to use one formulation. The problem has also been extended to incorporate upgrade considerations (Bagajewicz and Sánchez, 2000b) and maintenance costs (Sánchez and Bagajewicz, 2000). Bagajewicz (2000, 2001) reviews all these methods and also discusses the applications to bilinear and fully nonlinear systems.

The design of instrumentation networks for monitoring/parameter estimations can be summarized as follows:

Minimize {*Total Cost*}

subject to
- Estimability of Key Variables

Correspondence concerning this article should be addressed to M. Bagajewicz at bagajewicz@ou.edu.

- Precision of Key Variables
- Reliability of Key Variables
- Gross-Error Robustness

The total cost includes the maintenance cost, which regulates the availability of variables, a concept that substitutes reliability when the system is repairable (Bagajewicz, 2000).

Although the basic goal of monitoring systems is to provide a good estimate of the state of the system, alarm systems are designed to alert about process malfunction. In turn, process faults propagate throughout the process, altering the readings of instruments (pressures, temperatures, flow rates, and so on). Thus, these sensors should be able to determine departures from normal operation. In this sense, this task is different from that of gross error detection, which concentrates on instrument malfunction. As a consequence, the problem of designing an alarm system consists of determining the cost-optimal position of sensors, such that all process faults, single or multiple and simultaneous, can be detected and distinguished from instrument malfunction (biases). The first attempt to present a technique to locate sensors was done by Lambert (1977), who used fault trees based on failure probabilities. Failure probabilities are hard to assess, fault trees cannot handle cycles, and the construction of the tree is cumbersome for large-scale systems. Because of these limitations the technique has not been developed further.

For a fault-monitoring system to be useful in practice, it should be able to observe all the faults and identify their location and number to the maximum extent possible. This is called maximum fault resolution. Therefore, a *sensor network for maximum fault resolution is such that each fault has one and only one set of nodes from which it is observable.* Raghuraj et al. (1999) proposed a procedure to obtain such sets of nodes for maximum resolution. They also show that the sensor-location problem for multiple faults and maximum resolution can be solved as an extension of the single-fault assumption problem. The application of these algorithms to a continuously stirred tank reactor (CSTR) and a Fluid Catalytic Cracker (FCC) unit are presented by Raghuraj et al. (1999). They used directed graphs (DG), that is, graphs without signs. The arcs of the DG represent a "will cause" relationship, that is, an arc from node *A* to node *B* implies that *A* is a sufficient condition for *B*, which in general is not true for a signed DG, where an arc represents a "can cause" relationship. The strategy used to solve the problem is based on identifying directed paths from root nodes where faults can occur to nodes where effects can be measured, called the *observability set*. Of all these paths, the objective sought is to choose the minimal subset of sensors from the observability set that would have at least one directed path from every root node. Raghuraj et al. (1999) proposed a greedy search, which was later modified to remove redundant members of the observability set by a "backtracking algorithm." They do not guarantee optimality; that is, a set featuring the minimum number of sensors may not be found by this algorithm. Bhushan and Rengaswamy (2000a,b) extended these models to signed directed graphs and reliability maximization.

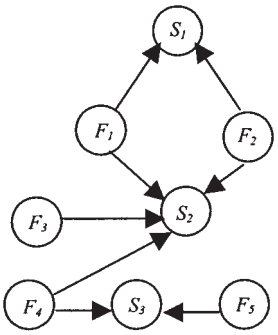

Figure 1. DG diagram of a hypothetical process.

The design procedure offered by Raghuraj et al. (1999) is not based on cost and does not include other features, such as filtering of sensor failures. However, very recently Bhushan and Rengaswamy (2001) presented a mixed-integer linear programming (MILP) formulation that maximizes reliability and uses cost constraints. Such an approach is equivalent to the minimization of cost subject to reliability thresholds (Bagajewicz and Sánchez, 2000a).

Thus, from the exclusive point of view of fault detection, the problem of the design of instrumentation is

Minimize {*Total Cost*}

subject to
- Desired Observability of Faults
- Desired Level of Resolution of Faults
- Desired level of Reliability of Fault Observation
- Desired level of Gross-Error Robustness in the Sensor Network

Gross-error robustness is the ability of the sensor network to filter undesired sensor failure. The combination of both goals—that is, the design of a sensor network capable of performing estimation of key variables for production accounting, parameter estimation, on-line optimization, and fault detection and diagnosis—is attempted in this report.

In this article, we specifically solve the following problem:

Minimize {*Total Cost*}

subject to
- Desired Observability of Faults
- Desired Level of Resolution of Faults
- Desired Precision level of the key variables.

We use an MILP formulation, part of which is taken from the MILP model developed by Bagajewicz and Cabrera (2001).

MILP Model for Fault Detection

Figure 1 depicts a digraph with the path connections from faults to potential location of sensors (DG diagram). In this figure, each node corresponding to a fault (F_i) connects through an arrow to a sensor (S_j), thus indicating that the fault will affect the reading of the corresponding sensor. Clearly, more than one sensor can be affected by a fault.

Let q be the vector of binary elements that denotes whether a sensor is installed ($q_i = 1$) or not installed ($q_i = 0$). With this definition the cost function becomes

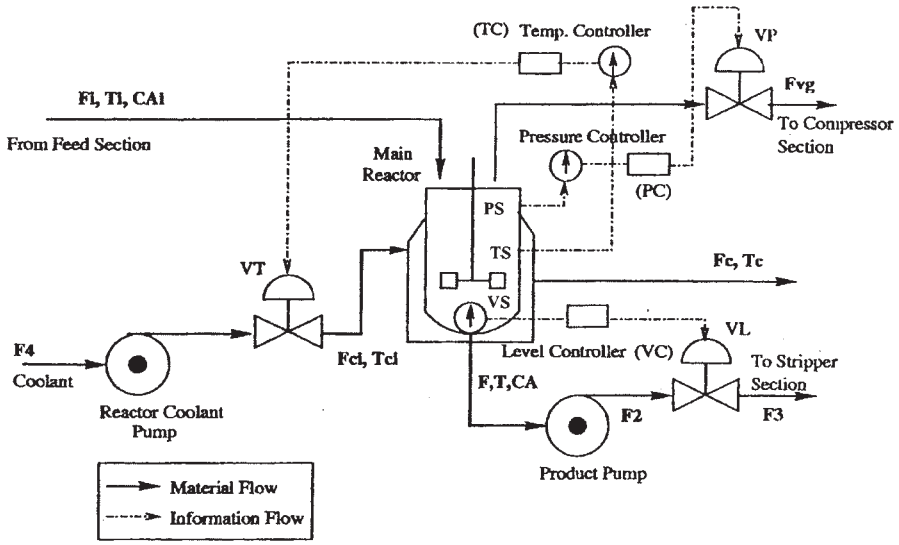

Figure 2. CSTR process.

$$\text{Min} \sum_{i \in M_1} c_i q_i \tag{1}$$

where M_1 is the set of potential sensor locations, that is, the dimension of q. The following step is to generate the fault-sensor maximum connectivity matrix (A), which will contain the same information as the bipartite graph. In this matrix, the columns represent the faults and the rows represent the sensors. When $A_{ij} = 1$, then we say that there is a path from fault F_j to the potential sensor location S_i. For the case of Figure 1, this matrix is the following

$$A = \begin{array}{ccccc|c} F_1 & F_2 & F_3 & F_4 & F_5 & \\ \hline 1 & 1 & 0 & 0 & 0 & S_1 \\ 1 & 1 & 1 & 1 & 0 & S_2 \\ 0 & 0 & 0 & 1 & 1 & S_3 \end{array} \tag{2}$$

We now define a fault-sensor connectivity matrix $B(q) = [b_1, \ldots, b_n]$ using

$$b_j = a_j \otimes q \tag{3}$$

where b_j and a_j are the ith column of B and A, respectively, and $\otimes$ is the Hadamard product of two vectors, that is, $(a_j \otimes q)_i = a_{ji} q_i$. Thus, $A = B$ only when sensors are selected for all potential locations. For the particular case of our example

$$B(q) = \begin{array}{ccccc|c} F_1 & F_2 & F_3 & F_4 & F_5 & \\ \hline q_1 & q_1 & 0 & 0 & 0 & S_1 \\ q_2 & q_2 & q_2 & q_2 & 0 & S_2 \\ 0 & 0 & 0 & q_3 & q_3 & S_3 \end{array} \tag{4}$$

We now observe that the sum of all the elements of one column of B determines whether the corresponding fault is either observable or not observable, that is

$$\text{If} \sum_{i=1}^{m} B_{ij} = 0 \rightarrow F_j \qquad \text{is unobservable} \tag{5}$$

$$\text{If} \sum_{i=1}^{m} B_{ij} \geq 1 \rightarrow F_j \qquad \text{is observable} \tag{6}$$

We are now in a position to present our model for cost-optimal sensor network design for fault observability

$$\left.\begin{array}{l} P1 = \text{Min} \sum_{i \in M_1} c_i q_i \\ \text{s.t.} \\ \sum_{i \in M_1} b_k(q_i) \geq 1 \qquad \forall\, k \in M_o \\ q_i \in \{0, 1\} \end{array}\right\} \tag{7}$$

where M_o is the set of faults that one wants to be able to observe and q_i is a binary variable as stated above. Bhushan and Rengaswami (2001) described this problem similarly and presented a greedy search algorithm to solve this problem. A special combination of faults to create "pseudo-faults" that would allow fault resolution was also presented, and an algorithm to remove redundant constraints was proposed (Bhushan and Rengaswami, 2000b). An alternative approach is used herein.

Fault Detection and Single-Fault Resolution

We now proceed to add fault resolution to the model. A fault can be observed because there are sensors located in the corresponding end of the path to candidate sensor location. However, the same type of symptoms can correspond to different

faults. Thus, the term "resolution" has been coined to refer to the ability of determining exactly which process fault occurred given the set of symptoms. Two faults can be distinguished only if one of them has at least one sensor that is different or does not sense the other fault at all. We now represent this using the connectivity matrix.

Consider two columns of the fault-sensor connectivity matrix. If both faults are observable, all one needs is these two faults not be connected to the same set of sensors. To guarantee that, at least one element of the two corresponding columns of B must be different. This is achieved by imposing the following condition

$$\sum_{i\in M_1} [b_j(q) \otimes b_p(q)]_i \le \text{Max}\left\{ \sum_{i\in M_1} [b_j(q)]_i, \sum_{i\in M_1} [b_p(q)]_i \right\} - 1 \tag{8}$$

In the matrix shown in Eq. 4, faults F_1 and F_4 are observable and distinguishable for certain set of sensors. Indeed, $\Sigma_{i\in M_1} [b_1(q) \otimes b_2(q)]_i = q_2^2$ and $\text{Max}\{\Sigma_{i\in M_1} [b_1(q)]_i, \Sigma_{i\in M_1} [b_4(q)]_i\} = \text{Max}\{q_1 + q_2, q_2 + q_3\}$. Thus, if $q_1 = 0$ and $q_3 = 0$, then the inequality expressed in Eq. 8 does not hold and the sensors cannot be distinguished (they would express both in sensor 2 only). In turn, if $q_2 = 0$, $q_1 = 1$, and $q_3 = 1$, then the inequality in Eq. 8 holds and the faults are distinguishable. They also are distinguishable; that is, the inequality in Eq. 8 holds if the three sensors are located ($q_1 = q_2 = q_3 = 1$). Now, consider faults F_1 and F_2. These two faults affect the same sensors and are indistinguishable, and there is no combination of sensors that can satisfy the inequality in Eq. 8. Indeed, in this case, $\Sigma_{i\in M_1} [b_1(q) \otimes b_2(q)]_i = q_1^2 + q_2^2$ and the sum of columns 1 and 2 are $q_1 + q_2$. Thus the right-hand side of Eq. 8, that is $q_1 + q_2 - 1$, which cannot be larger than the left-hand side, no matter what values assume q_1 and q_2. This result also suggests that one cannot incorporate Eq. 8 for these two sensors in the model because it will become infeasible.

Thus the final model for fault observability and single fault resolution is

$$\left.\begin{array}{l} P2 = \text{Min} \sum_{i\in M_1} c_i q_i \\ \text{s.t.} \\ \sum_{i\in M_1} a_{ki} q_i \ge 1 \qquad \forall\, k \in M_O \\ \sum_{i\in M_1} a_{ji} a_{pi} q_i \le C_{jp} - 1 \qquad \forall\, j, p \in M_S; j \ne p \\ C_{jp} = \text{Max}\left\{ \sum_{i\in M_1} a_{ji} q_i, \sum_{i\in M_1} a_{pi} q_i \right\} \qquad \forall\, j, p \in M_S; j \ne p \\ q_j \in \{0, 1\} \qquad \forall\, j \in M_S \end{array}\right\} \tag{9}$$

where M_S is the set of faults that one wants to be able to observe and resolve. Note that M_S should be included in M_O.

The Max operator in Eq. 8 is not amenable to be used in MILP models. To convert it to a linear set of inequalities, consider the following equivalency

$$C = \text{Max}(A, B) \Leftrightarrow \begin{cases} C - \boldsymbol{\Omega} \cdot K \le A & C \ge A \\ C - (1 - K) \cdot \boldsymbol{\Omega} \le B & C \ge B \end{cases} \tag{10}$$

where K is binary and $\boldsymbol{\Omega}$ is an upper bound on the total number of sensors. The equivalency established in Eq. 10 forces the equality of C to the largest of A and B, by selecting $K = 1$ or $K = 0$.

Fault Detection and Multiple-Fault Resolution

Assume now that one wants to consider multiple faults occurring simultaneously and one wants to resolve them. Given that multiple faults can have the same signature as that of single faults, one wants to resolve multiple faults from each other, regardless of their cardinality. That is, if one targets as a design procedure that one desires resolution of two faults at a time, then one needs to be able to distinguish them not only from each other but also from single faults. We now illustrate how this can be done for two faults.

Consider two faults taking place at the same time. They will have a signature corresponding to a vector that is the Boolean sum of the corresponding two columns of B, that is, their signature will be

$$d_{jk}(q) = [a_j \oplus a_k] \otimes q \tag{11}$$

Thus, we define the double fault-sensor connectivity matrix $D(q)$, through all column vectors, as the desired combinations of double faults $d_{jk}(q)$ that belong to a prespecified set. Resolution of these faults with respect to the others is now obtained through an expression similar to Eq. 8, that is

$$\sum_{i\in M_1} [d_{jk}(q) \otimes d_{pr}(q)]_i \le \text{Max}\left\{ \sum_{i\in M_1} [d_{jk}(q)]_i, \sum_{i\in M_1} [d_{pr}(q)]_i \right\} - 1 \qquad \forall\, j, k, p, r \in M_{S2} \tag{12}$$

where (j, k) and (p, r) belong to M_{S2}, the set of all combinations of two faults that one wants to resolve. In addition $j \ne k \ne p \ne r$. This equation can be replaced by a set of mixed integer linear equations using Eq. 10. In a similar fashion, one can distinguish double faults from single faults by using the following expression

$$\sum_{i\in M_1} [d_{jk}(q) \otimes b_p(q)]_i \le \text{Max}\left\{ \sum_{i\in M_1} [d_{jk}(q)]_i, \sum_{i\in M_1} [b_p(q)]_i \right\} - 1 \qquad \forall\, j, k \in M_{S2},\ \forall\, p \in M_S \tag{13}$$

where $j \neq k \neq p$ and which can also be rewritten as a set of mixed-integer linear equations using Eq. 10. The final model for observability, single- and double-fault resolution, is

$$
\begin{aligned}
&P3 = \text{Min} \sum_{i \in M_1} c_i q_i \\
&\text{s.t.} \\
&\sum_{i \in M_1} a_{ki} q_i \geq 1 \qquad \forall\, k \in M_O \\
&\sum_{i \in M_1} a_{ji} a_{pi} q_i \leq C_{jp} - 1 \qquad \forall\, j, p \in M_S; j \neq p \\
&C_{jp} = \text{Max}\left\{ \sum_{i \in M_1} a_{ji} q_i, \sum_{i \in M_1} a_{pi} q_i \right\} \qquad \forall\, j, p \in M_S; j \neq p \\
&\sum_{i \in M_1} [(a_{ji} \oplus a_{ni}) \cdot (a_{pi} \oplus a_{ri}) \cdot q_i] \leq G_{jnpr} \qquad \forall\, j, n, p, r \in M_{S2}; j \neq n \neq p \neq r \\
&G_{jnpr} = \text{Max}\left\{ \sum_{i \in M_1} [(a_{ji} \oplus a_{ni}) q_i], \sum_{i \in M_1} [(a_{pi} \oplus a_{ri}) q_i] \right\} \qquad \forall\, j, n, p \in M_S; j \neq n \neq p \neq r \\
&\sum_{i \in M_1} [(a_{ji} \oplus a_{ni}) \cdot a_{pi} \cdot q_i] \leq Q_{jnp} \qquad \forall\, j, n \in M_{S2},\ \forall\, p \in M_S; j \neq n \neq p \\
&Q_{jnp} = \text{Max}\left\{ \sum_{i \in M_1} [(a_{ji} \oplus a_{ni}) q_i], \sum_{i \in M_1} [a_{pi} q_i] \right\} \qquad \forall\, j, p \in M_S; j \neq n \neq p
\end{aligned}
\tag{14}
$$

where M_S and M_{S2} are the sets of faults that one wants to be able to observe and resolve individually and simultaneously and M_O is the set of faults that one wants to be able to simply observe, respectively. Note, that M_S as well as M_{S2} should be included in M_O.

Our methodology, requires the identification of pseudofaults (Eq. 10) for the double-fault resolution problem, but this can be done automatically. An alternative methodology was presented by Bushan and Rengaswamy (2001). In their work, the number of constraints for the resolution problem is four times lower, but identification of pseudofaults is required for any resolution problem considered.

MILP Model for Variable Precision

We now briefly review the model developed by Bagajewicz and Cabrera (2001) for the design of instrumentation networks, with the objective of regular monitoring/parameter estimation. The model includes several constraints (precision, residual precision, error detectability, and resilience).

The cost-optimal sensor network of a system with accuracy constraints is obtained by solving the following optimization problem (Bagajewicz, 1997)

$$
\left.
\begin{aligned}
&\text{Min} \sum_i c_i q_i \\
&\text{s.t.} \\
&\hat{s}_{ii} < s_{ii}^* \qquad \forall\, i \in M_P
\end{aligned}
\right\}
\tag{15}
$$

where q_i is the binary vector indicating that a sensor standard deviation s_{ii} is located in variable i, c_i is the cost of such sensor. The precision constraint states that the variance of estimates $\hat{s}_{ii}$ (after reconciliation) has to be strictly smaller than a certain threshold s_{ii}^*. The constraint is written only for the set of variables of interest (M_P). When the equations of the system correspond to material balances, they are linear and they can be expressed by $A \cdot x = 0$, where A is the matrix of coefficients and x represents the mass flows. The variance of the estimates of measured and redundant variables obtained after data reconciliation is performed is given by

$$
\hat{S}_R = S_R - S_R \cdot E_R^T \cdot (E_R \cdot S \cdot E_R^T)^{-1} \cdot E_R \cdot S_R \tag{16}
$$

where E_R is the matrix reflecting the material balance equations of measured and redundant variables. The diagonal elements of $\hat{S}_R$ are the values of interest $\hat{s}_{ii}$, whereas the diagonal elements of S_R are the variances of the sensors installed in the redundant variables.

The MILP model developed by Bagajewicz and Cabrera (2001) considers that all variables are measured, some with real instruments and others with fake instruments of very high standard deviation, high enough so that they do not have an influence in the other variables after data reconciliation. The model also includes hardware redundancy, and considers many sensor candidates for each variable. When such an approach is used, $E_R = A$, and $S_R = S$. The complete model without hardware redundancy is

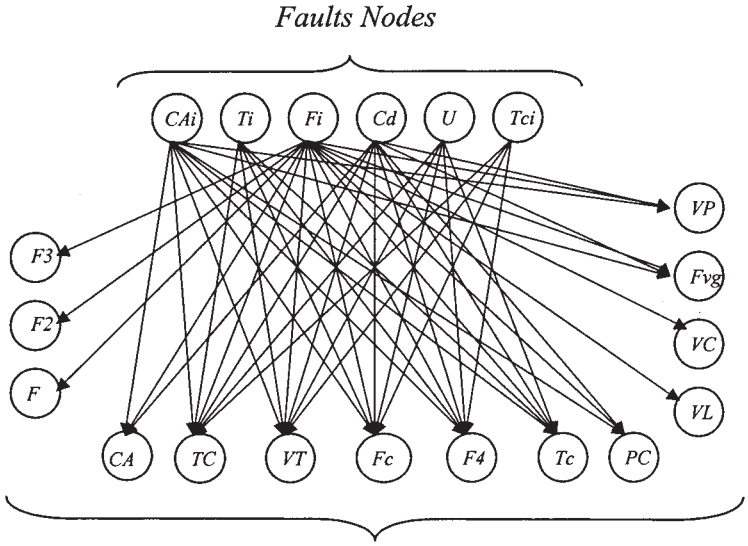

Figure 3. DG diagram for noncontrol-loop faults.

$$\left.\begin{aligned}
&\text{Min} \sum_{i\in M_1}\sum_{j=1}^{l_i^m} c_{ij}\mu_{ij}\\
&\text{s.t.}\\
&\sum_{j=1}^{l_i^m}\mu_{ij}=1 \qquad \forall\, i\in M_1\\
&\sum_{t=1}^{l_i^m}\mu_{it}\sigma_{it}^2-\sum_{t=1}^{l_i^m}\sum_{k=1}^{n}\sum_{d=1}^{n}\mu_{it}\sigma_{it}^4 a_{ki}p_{kd}a_{di}\le \hat{s}_{ii} \qquad \forall\, i\in M_p\\
&\sum_{h=1}^{n}\sum_{k=1}^{m}\sum_{t=1}^{l_i^m} p_{ih}a_{hk}\sigma_{kt}^2\mu_{kt}a_{jk}=\delta_{ij} \qquad \forall\, i\in M_1, j=1,\ldots,n
\end{aligned}\right\} \tag{17}$$

where μ_{it} is a binary variable indicating that a sensor standard deviation s_{it} is located in variable i, and σ_{ii}^2 are the variances of sensor t installed in variable i. The last two equations of the above model represent Eq. 16. They are the result of introducing a new matrix as follows: $P = (A \cdot S \cdot A^T)^{-1}$, which is summarized in the last equation. Thus, a_{ki} are elements of matrix A, p_{kd} are the elements of P, and δ_{ij} is the identity matrix. Finally, l_i^m is the number of different alternative candidates of measurement devices for variable i, n is the number of material balance equations, and m is the number of variables. The model contains products of integers and continuous variables (p_{kd} and μ_{it}), which are further linearized using a transformation proposed by Glover (1975).

This model can be extended to nonlinear systems by simple linearization. Indeed, when $F(x) = 0$ is the nonlinear system of equations representing the model, using a Taylor series expansion that ignores terms of order > 1 gives

$$\left[\frac{\partial F(x)}{\partial x}\right]_{x=x_0} x = F(x_0) - \left[\frac{\partial F(x)}{\partial x}\right]_{x=x_0} x_0 \tag{18}$$

Now, defining $z = x - \alpha$, where α is such that

$$\left[\frac{\partial F(x)}{\partial x}\right]_{x=x_0} \alpha = F(x_0) - \left[\frac{\partial F(x)}{\partial x}\right]_{x=x_0} x_0 \tag{19}$$

The system of Eq. 18 becomes

$$\left[\frac{\partial F(x)}{\partial x}\right]_{x=x_0} z = 0 \tag{20}$$

which has the required form. To avoid numerical problems the new variables (z) are scaled so that $\text{Max}_{\forall j}|A_{ij}| = 1$, $\forall\, i$.

Case Study

We considered the CSTR example introduced by Bhushan et al. (2000) (Figure 2). The equations were implemented in GAMS and solved using CPLEX. First, the problem of fault detection/resolution and the precision-constrained problem were solved separately. Next, all constraints are used simultaneously to show how the synergism between the two constraints forces solutions different from the simple superposition of separate solutions.

The CSTR model equation and the correspondent DG are presented in the Appendix. Following the strategy of Bhushan et al. (2000), the corresponding sensor nodes for each fault were identified. Figure 3 shows the bipartite graph for observability when control-loop faults are excluded. Table A2 can be used to generate the graphs including control-loop faults.

The model was first run minimizing the number of sensors (that is, using the same cost for all sensors). Sensors *TS*, *PS*, and *VS* are installed for control purposes and can be used for fault detection. They are assigned zero cost. When only observability and single- as well as double-fault resolution are requested (but no precision on variables is imposed), the results

Table 1. Minimum Number of Sensors Solutions for the CSTR System

	Case 1						Case 2	
Node Name	Cost	q Alternative				Bushan et al.	Cost	q
CA	100	X				X	100	
TS	0	X	X	X	X	X	0	X
TC	100						100	
VT	100						100	
F_c	100	X	X	X	X	X	1	X
F_4	100						1	
T_c	100	X	X	X	X	X	1	X
N	100						—	
PS	0	X	X	X	X	X	0	X
PC	100			X		X	100	
VP	100				X		100	
F_{vg}	100		X				1	X
VS	0	X	X	X	X	X	0	X
VC	100						100	
VL	100						100	
F_3	100						1	
F_2	100						1	
F	100	X	X	X	X	X	1	X
Number of sensors		7				8		7

obtained are shown in Table 1, Case 1. A cross symbol ($\times$) indicates that the corresponding node should be measured. Table 1 also shows that an alternative arrangement of sensors can be obtained using the same cost for all measurements. These four solutions are different from the one obtained by Bhushan et al. (2000), which is also shown in Table 1. We also modified the problem costs (Case 2). For this last case we consider that the cost of measuring concentration is too high and the node N (number of mol in gaseous phase) is unmeasurable, thus precluding it. Furthermore, we assign high costs to the use of controllers and valves for detecting faults.

A few observations are appropriate:

- For this small example, we found four alternative solutions that feature a minimum number of sensors (seven).
- If costs are considered, fewer alternative solutions featuring the same cost are likely, but one could expect to identify suboptimal solutions with costs close to the optimal.
- Some nodes are essential and will be part of all solutions, no matter what cost is assigned to measure them. This is the case of T_c, F_c, and F. In the work of Bushan et al. (2000) node F_c is the node where the sensor should be placed to observe all the faults (except control-loop failures) under the single-fault assumption. For observability under the double-fault assumption (control loop–noncontrol-loop faults and noncontrol–noncontrol faults), node T_c has been selected to observe such pairs. Finally, node F is the node where the sensor should be placed for solving the single-fault resolution problem.
- Those faults that affect the same set of nodes are not distinguished.

Next, the problem was run requiring only precision in key variables ($\sigma_{CA} = 2\%$, $\sigma_T = 1\%$, and $\sigma_F = 2\%$). For each variable from the precision group (F_i, C_{Ai}, C_A, T, T_i, T_c, F_c, T_{ci}, F_{vg}, F, F_2, F_3, F_4) sensors of $\sigma_1 = 1\%$ are available at a cost of 100. Consider also that for each node not considered in the precision problem the cost of locating a sensor is 100 (as in Case 1). The optimal set of sensors for this problem is shown in Table 2. The results for higher precision are shown in Table 3.

Consider first the case of Table 2. Although 7 sensors are enough for fault detection, the same number of sensors is sufficient for fault detection and the estimation of the desired variables, although the sensors are different. Note that the simple superposition of solutions for fault detection and precision requires 7 and 10 sensors as minimum and maximum, respectively. Table 3 illustrates that this synergism does not always reduce the number of sensors. Indeed, 9 sensors are required to achieve both goals, whereas the superposition of solutions requires 10 and 11 sensors as minimum and maximum, respectively.

Conclusions

An MILP formulation for the location of sensors for fault detection was presented. The technique provides a very useful tool to design and upgrade instrumentation networks for fault observability and resolution. This formulation of the problem allows the introduction of cost as one more element of consideration in the design of these systems.

Results show that the model can be used to identify alternative solutions. The example chosen is very revealing because it shows that many solutions with the same fault observability and resolution properties feature the same low number of sensors. The use of the model also reveals that one can identify solutions with a larger number of sensors, which could be of interest for other reasons. Extensions of this work consist of the addition of reliability and robustness constraints, the addition of residual precision, and gross error robustness goals used for monitoring of normal operations and the treatment of this problem as a multiobjective case.

Acknowledgments

Financial support from the National Science Foundation through Grant CTS-0075814 and the University of Oklahoma for matching funds is gratefully acknowledged.

Parts of this article were presented at the 4th IFAC Workshop on On-Line Fault Detection and Supervision in the Chemical Process Industries (Bagajewicz and Fuxman, 2001).

Notation

A = maximum connectivity matrix

Table 2. Precision Required for *CA*, *T*, *F

Node	Sensors for Fault Detection				Sensors for Precision (A)												Sensors for Fault Detection and Precision (B)	Precision after Reconciliation (B)
F_i					x	x	x											1%
C_{Ai}					x			x	x	x								13.33%
C_A	x				x	x	x	x	x	x	x	x	x	x	x	x	x	1%
TS	x	x	x	x		x					x	x	x				x	1%
T_i																		18.56%
T_c	x	x	x	x													x	1%
F_c	x	x	x	x													x	1%
T_{ci}																		10.94%
F_{vg}		x					x							x	x	x		25.05%
F	x	x	x	x				x			x			x			x	1%
F_2									x			x			x			1%
F_3										x			x			x		1%
F_4																		1%
TC																		
VT																		
N																		
PS	x	x	x	x													x	
VS	x	x	x	x													x	
PC			x															
VP																		
VC																		
VL																		

*Precision threshold $\sigma^*_{CA} \leq 2\%$, $\sigma^*_T \leq 1\%$, $\sigma^*_F \leq 2\%$.

a_j = column of matrix A
$B(q)$ = fault-sensor connectivity matrix
b_j = column of matrix $B(q)$
c_i = cost of sensor i
$D(q)$ = double-fault–sensor connectivity matrix
d_{jk} = column of matrix $D(q)$
K = binary variable
M_O = set of faults that one wants to be able to observe
M_S = set of faults that one wants to be able to observe and resolve
M_{S2} = set of all combinations of two faults that one wants to resolve
PS = pressure sensor
q_i = binary variable indicating weather a sensor is located (1) or not (0)
R_i = fault i
S_i = sensor that corresponds to variable i

Greek letter

Ω = upper bound (total number of sensors)

Table 3. Higher Precision Required for *CA*, *T*, *F

Node	Sensors for Fault Detection				Sensors for Precision (A)		Sensors for Fault Detection and Precision (B)	Precision After Reconciliation (A)/(B)
F_i								0.923%, 0.923%/0.923%
C_{Ai}					x	x	x	0.658%, 0.658%/0.658%
C_A	x				x	x	x	0.935%, 0.935%/0.935%
TS	x	x	x	x			x	0.095%, 0.134%/0.056%
T_i								—, —/16.729%
T_c	x	x	x	x			x	—, —/1%
F_c	x	x	x	x			x	—, —/1%
T_{ci}								—, —/8.23%
F_{vg}		x			x	x	x	0.917%, 0.917%/0.916%
F	x	x	x	x			x	0.923%, 0.923%/0.923%
F_2					x			0.923%, 0.923%/0.923%
F_3						x		0.923%, 0.923%/0.923%
F_4								—, —/1%
TC								
VT								
N								
PS	x	x	x	x			x	
VS	x	x	x	x			x	
PC			x					
VP								
VC								
VL								

*Precision threshold $\sigma^*_{CA} = \sigma^*_T = \sigma^*_F \leq 0.95\%$.

Literature Cited

Alhéritière, C., N. Thornhill, S. Fraser, and M. Knight, "Evaluation of the Contribution of Refinery Process Data to Performance Measures," AIChE Annual Meeting, Los Angeles, CA (1997).

Alhéritière, C., N. Thornhill, S. Fraser, and M. Knight, "Cost Benefit Analysis of Refinery Process Data: Case Study," *Comp. Chem. Eng.*, **22**(Suppl.), S1031 (1998a).

Alhéritière, C., N. Thornhill, S. Fraser, and M. Knight, "Cost Benefit Analysis of Process Data in Plant Performance Analysis," AIChE Annual Meeting, Miami, FL (1998b).

Ali, Y., and S. Narasimhan, "Sensor Network Design for Maximizing Reliability of Linear Processes," *AIChE J.*, **39**(5), 2237 (1993).

Ali, Y., and S. Narasimhan, "Sensor Network Design for Maximizing Reliability of Bilinear Processes," *AIChE J.*, **42**(9), 2563 (1996).

Bagajewicz, M., "Optimal Sensor Location in Process Plants," *AIChE J.*, **43**(9), 2300 (1997).

Bagajewicz, M., *Design and Upgrade of Process Plant Instrumentation*, Technomic Publishers, Lancaster, PA (2000).

Bagajewicz, M., "Review of Recent Results in Instrumentation Design and Upgrade for Process Plants," Proc. of the 4th IFAC Workshop on On-Line Fault Detection & Supervision in the Chemical Process Industries, June 8–9, Seoul, Korea (2001).

Bagajewicz, M., and E. Cabrera, "A New MILP Formulation for Instrumentation Network Design and Upgrade," Proc. of the 4th IFAC Workshop on On-Line Fault Detection and Supervision in the Chemical Process Industries, June 8–9, Seoul, Korea (2001a).

Bagajewicz, M., and E. Cabrera, "A Multiobjective Approach for Instrumentation Network Design and Upgrade," Proc. of Enpromer 2001, 3rd Mercosur Congress on Process Systems Engineering, Sep. 16–20, Santa Fe, Argentina (2001b).

Bagajewicz, M., and E. Cabrera, "A New MILP Formulation for Instrumentation Network Design and Upgrade," *AIChE J.*, **48**(10), 2271 (2002).

Bagajewicz, M., and A. Fuxman, "An MILP Model for Cost Optimal Instrumentation Network Design and Upgrade for Fault Detection," Proc. of the 4th IFAC Workshop on On-Line Fault Detection and Supervision in the Chemical Process Industries, June 8–9, Seoul, Korea (2001).

Bagajewicz, M., and M. Sánchez, "Duality of Sensor Network Design Models for Parameter Estimation," *AIChE J.*, **45**(3), 661 (1999a).

Bagajewicz, M., and M. Sánchez, "Design and Upgrade of Non-Redundant and Redundant Linear Sensor Networks," *AIChE J.*, **45**(9), 1927 (1999b).

Bagajewicz, M., and M. Sánchez, "Cost-Optimal Design of Reliable Sensor Networks," *Comput. Chem. Eng.*, **23**(11/12), 1757 (2000a).

Bagajewicz, M., and M. Sánchez, "Reallocation and Upgrade of Instrumentation in Process Plants," *Comput. Chem. Eng.*, **24**(8), 1961 (2000b).

Bhushan, M., and R. Rengaswamy, "Design of Sensor Location Based on Various Fault Diagnosis Observability and Reliability Criteria," *Comput. Chem. Eng.*, **24**, 735 (2000a).

Bhushan, M., and R. Rengaswamy, "Design of Sensor Network Based on the Signed Directed Graph of the Process for Efficient Fault Diagnosis," *Ind. Eng. Chem. Res.*, **39**, 999 (2000b).

Bhushan, M., and R. Rengaswamy, "A Framework for Sensor Network Design for Efficient and Reliable Fault Diagnosis," Proc. of the 4th IFAC Workshop on On-Line Fault Detection and Supervision in the Chemical Process Industries, June 8–9, Seoul, Korea (2001).

Glover, F., "Improved Linear Integer Programming Formulations of Nonlinear Integer Problems," *Man. Sci.*, **22**(4), 455 (1975).

Lambert, H. E., "Fault Trees for Locating Sensors in Process Systems," *CEP*, **Aug.**, 81 (1977).

Luong, M., D. Maquin, C. T. Huynh, and J. Ragot, "Observability, Redundancy, Reliability and Integrated Design of Measurement Systems," Proc. of 2nd IFAC Symposium on Intelligent Components and Instrument Control Applications, Budapest (1994).

Madron, F., *Process Plant Performance, Measurement Data Processing for Optimization and Retrofits*, Ellis Horwood, West Sussex, UK (1992).

Madron, F., and V. Veverka, "Optimal Selection of Measuring Points in Complex Plants by Linear Models," *AIChE J.*, **38**(2), 227 (1992).

Maquin, D., M. Luong, and J. Ragot, "Observability Analysis and Sensor Placement," SAFE PROCESS '94 IFAC/IMACS Symposium on Fault Detection, Supervision and Safety for technical Process, June 13–15, Espoo, Finland (1994).

Narasimhan, S., and C. Jordache, *Data Reconciliation and Gross Error Detection. An Intelligent Use of Process Data*, Gulf Publishing Company, Houston, TX (2000).

Raghuraj, R., M. Bhushan, and R. Rengaswamy, "Locating Sensors in Complex Chemical Plants Based on Fault Diagnostic Observability Criteria," *AIChE J.*, **45**(2), 310 (1999).

Sánchez, M., and M. Bagajewicz, "On the Impact of Corrective Maintenance in the Design of Sensor Networks," *Ind. Eng. Chem. Res.*, **39**(4), 977 (2000).

Sen, S., S. Narasimhan, and K. Deb, "Sensor Network Design of Linear Processes using Genetic Algorithms," *Comput. Chem. Eng.*, **22**(3), 385 (1998).

Viswanath A., and S. Narasimhan, "Multi-Obective Sensor Network Design Using Genetic Algorithms," Proc. of 4th IFAC Workshop on On-Line Fault Detection & Supervision in the Chemical Process Industries, June 8–9, Seoul, Korea (2001).

Appendix

The CSTR given by Bushan at al. (2000a), involves an exothermic liquid-phase reaction: $A_{(l)} \rightarrow B_{(l)} + C_{(g)}$. The temperature controller controls the temperature of the reactor by manipulating the flow rate of the coolant flowing through the jacket. The level in the reactor is controlled by the level controlled by manipulating the outlet flow rate from the reactor. The pressure in the reactor is controlled by changing the vent gas flow rate. Both the reactor and the jacket are modeled with perfectly mixed-tank dynamics. The model equations are as follows.

Global Mass Balance

$$F_i - F = \frac{dV}{dt} \tag{A1}$$

Component Mass Balance (C_A)

$$\frac{F_i}{V}(C_{Ai} - C_A) - r_A = \frac{dC_A}{dt} \tag{A2}$$

Overall Heat Balance on the Reactor. Result obtained assuming constant heat capacities and densities:

$$\frac{F_i}{V}(T_i - T) + \frac{r_A(-\Delta H)}{\rho C_p} - \frac{UA(T - T_c)}{V\rho C_p} = \frac{dT}{dt} \tag{A3}$$

Overall Heat Balance on the Jacket

$$\frac{F_c}{V_j}(T_{ci} - T_c) + \frac{UA(T - T_c)}{V_j\rho_j C_{pj}} = \frac{dT_c}{dt} \tag{A4}$$

Gas Phase Balance

$$r_A V - F_{vg} = \frac{dn}{dt} \tag{A5}$$

Reaction Rate

$$r_A = C_d C_A k_0 e^{-E/RT} \tag{A6}$$

Table A1. Corresponding Nodes for Noncontrol Loop Faults

	C_{Ai}	F_i	T_i	T_{ci}	C_d	U
C_A	x	x			x	
T_c	x	x	x		x	x
F_c	x	x	x	x	x	x
F_{vg}	x	x			x	
F		x				
F_2		x				
F_3		x				
F_4	x	x	x	x	x	x
TC	x	x	x	x	x	x
VT	x	x	x	x	x	x
N						
PC	x	x			x	
VP	x	x			x	
VC		x				
VL		x				

Table A3. Values for the CSTR of Figure 2

Notation	Variable	(Steady State/Constant) Value
V	Volume of reactor	48 ft^3
C_A	Reactant concentration in reactor	0.2345 lb · mol of A/ft^3
T	Reactor temperature	600°R
F	Outlet flow rate	40 ft^3/h
N	No. of moles of vapor	28.3657 lb · mol
P	Pressure in vapor space	2116.79 lb/ft^2
F_{vg}	Vent flow rate	10.6137 lb · mol/h
F_i	Inlet feed flow rate	40 ft^3/h
C_{Ai}	Inlet reactant concentration	0.5 lb · mol of A/ft^3
T_c	Jacket temperature	590.51°R
F_c	Coolant flow rate	56.626 ft^3/h
T_i	Inlet feed temperature	530°R
V_j	Volume of jacket	3.85 ft^3
k_0	Frequency factor	$7.08 \times 10^{10}\ h^{-1}$
C_d	Catalyst activity	1
E	Activation energy	29900 btu/lb · mol
R	Universal gas constant	1.99 btu/lb · mol °R
U	Heat-transfer coefficient	150 btu/h ft^2 °R
A	Heat-transfer area	150 ft^2
T_{ci}	Inlet coolant temperature	530°R
ΔH	Heat of reaction	−30,000 btu/lb · mol
C_p	Heat capacity (process side)	0.75 btu/lb_m °R
C_{pj}	Heat capacity (coolant side)	1 btu/lb_m °R
ρ	Density of process mixture	50 lb_m/ft^3
ρ_j	Density of coolant	62.3 lb_m/ft^3
V_g	Volume of vapor space	16 ft^3

Table A2. Corresponding Nodes for Control Loop Faults (Valve Failures)

	CA_i– VfL	CA_i– VfP	CA_i– VfT	T_i– VfL	T_i– VfP	T_i– VfT	F_i– VfL	F_i– VfP	F_i– VfT	T_{ci}– VfL	T_{ci}– VfP	T_{ci}– VfT	C_d– VfL	C_d– VfP	C_d– VfT	U– VfL	U– VfP	U– VfT
C_A	x	x	x			x	x	x	x			x	x	x	x			x
T_c	x	x	x		x	x	x	x	x			x	x	x	x	x	x	x
F_c	x	x		x	x		x	x		x	x		x	x		x	x	
F_{vg}	x	x	x			x	x	x	x			x	x	x	x			x
F								x	x									
F_2								x	x									
F_3								x	x									
F_4	x	x		x	x		x	x		x	x		x	x		x	x	
TC	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
VT	x	x		x	x		x	x		x	x		x	x		x	x	
N		x	x			x		x	x			x	x	x	x			x
PC	x	x	x			x	x	x	x			x	x	x	x			x
VP	x		x			x	x		x			x	x		x			x
VC							x	x	x									
VL								x	x									
TS			x			x			x			x			x			x
PS		x						x			x			x				
VS							x											

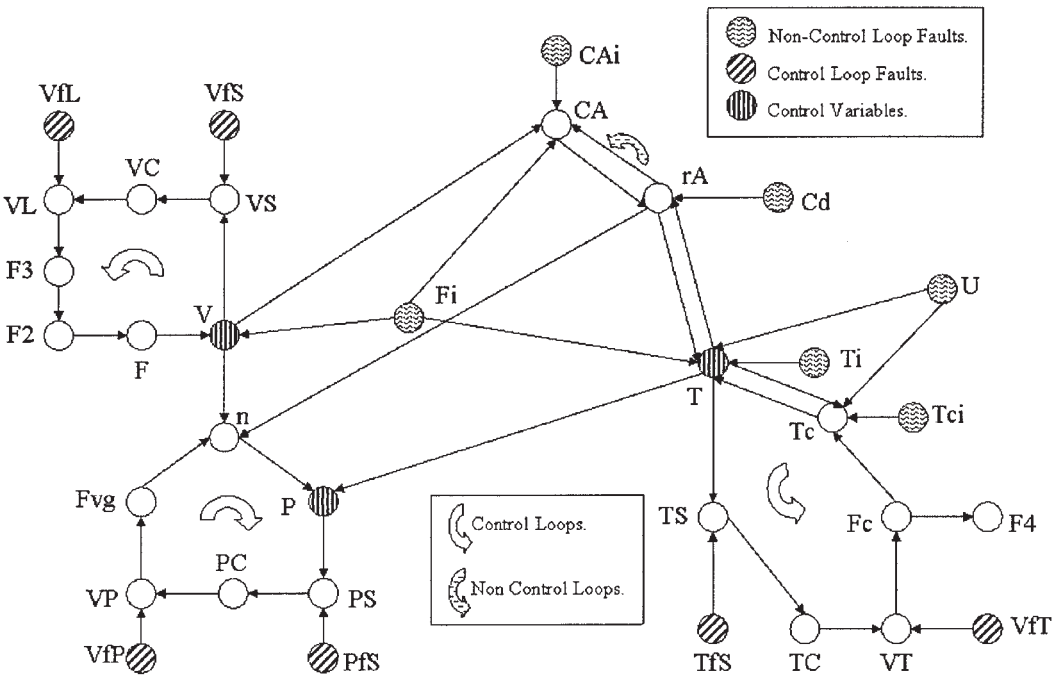

Figure A1. DG diagram of the CSTR.

Elemental Mass Balances in Valves and Pumps. Assuming no accumulation:

$$\left.\begin{aligned} F_3 - F_2 = 0 \\ F_2 - F = 0 \\ F_4 - F_c = 0 \end{aligned}\right\} \tag{A7}$$

Pressure in the Reactor. Assuming ideal behavior:

$$PV_g = nRT \tag{A8}$$

where V_g is the vapor space and is assumed constant.

Using Eqs. A1 through A8, the DG diagram (Figure A1) is generated. In this diagram, *PfS*, *TfS*, and *VfS* correspond to sensor fault nodes and *VfP*, *VfT*, and *VfL*, to valve fault nodes.

Table A1 shows the nodes for faults that do not correspond to the control loop. Table A2 shows the nodes for faults in the control loop (valve faults) associated with changes in C_{Ai}, F_i, T_i, T_{ci}, C_d, and U. The corresponding nodes for sensors failures are omitted. For the single-fault resolution problem, faults in Table A1 are considered. For the double-fault resolution, faults in Tables A1 and A2 are used.

The Jacobian (matrix E_R in Eq. 16) is the following:

$$\left[\frac{\partial F(x)}{\partial x}\right]_{x=x_0} = \begin{bmatrix} F_i & C_{Ai} & C_A & T & T_i & T_c & F_c & T_{ci} & F_{vg} & F & F_2 & F_3 & F_4 \\ -5.312\text{E-}3 & -0.8333 & 1.7763 & 9.2287\text{E-}3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1.4583 & 0 & -75440 & 5.9503 & -0.8333 & -125 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -93.8067 & 0 & 1085147 & 15.7169 & -14.708 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -45.2612 & -0.443 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Note that V, V_g, U, A, C_d, densities, and heat capacities are considered constants. The values of the variables used in the equations are shown in Table A3.

Manuscript received Aug. 7, 2001, and revision received Dec. 1, 2003.