Development and Implementation of a High-Performance Sensor System for an Industrial Polymer Reactor

ARTICLE *in* INDUSTRIAL & ENGINEERING CHEMISTRY RESEARCH · APRIL 2005

Impact Factor: 2.59 · DOI: 10.1021/ie049614t

3 AUTHORS, INCLUDING:

Babatunde A Ogunnaike
University of Delaware
189 PUBLICATIONS **2,268** CITATIONS

SEE PROFILE

James Schwaber
Thomas Jefferson University
146 PUBLICATIONS **3,817** CITATIONS

SEE PROFILE

Development and Implementation of a High-Performance Sensor System for an Industrial Polymer Reactor

Martin Pottmann

E.I. DuPont de Nemours and Company, Wilmington, Delaware 19880

Babatunde A. Ogunnaike*

Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716

James S. Schwaber

Department of Pathology, Anatomy and Cell Biology, Thomas Jefferson University, Philadelphia, Pennsylvania 19107

Because of increasingly stringent demands on product quality, process measurements of high precision, accuracy, and reliability are required in many processes where process variables are traditionally difficult to measure. In many applications, several different measurements of certain key process variables are available from various sources, including on-line analyzers, lab analyses, inferential measurements, and models of all types. Because these individual sensors often differ significantly in terms of their static and dynamic characteristics, precision, accuracy, sampling rate, measurement time delay, etc., in practice, only one of them is utilized for process monitoring and control purposes; the other sensor measurements are ignored. However, significantly improved estimates of the process variables can be obtained by utilizing all of the available sensor information simultaneously, i.e., by combining (fusing) these dissimilar measurements into a single estimate in a robust and fault-tolerant fashion. There is evidence that such sensor fusion operations are integral to the operation of biological control systems, and it is from one of these systems (the blood pressure control system) that we have drawn inspiration for developing the technique discussed in this article. A sensor fusion approach is developed using stochastic systems theory (in particular, Kalman filtering). The advantages of using multiple, redundant sensors, as well as the potential improvements achievable by combining delayed or multirate sensor information with that of a single nominal sensor, are then quantified. Robustness is achieved by augmenting the nominal technique with failure detection, classification, and compensation schemes, based on statistical hypothesis testing. The performance of the techniques is demonstrated on an actual plant application in which improved feed composition estimates are obtained for an industrial ethylene copolymerization reactor.

1. Introduction

In many processes, the main barrier to obtaining high-quality measurements is the fact that many key process variables are difficult to measure online. Reliable measurements can only be obtained by augmenting online sensors with laboratory analyses, or by inferential deduction from auxiliary measurements. Therefore, techniques are needed that are capable of combining information about a process variable from a variety of sources, including online instruments, offline analyzers, and models of all types (e.g., fundamental models, artificial neural networks, regression models). The specific *sensor fusion* problem addressed in this article can be stated as follows: "*Given measurements of a single process variable that are available from different sources, at possibly different rates, and with different measurement delays, form a composite measurement in a robust and failure tolerant fashion.*"

Many practical examples exist in which measurements of a key process variable are available from several different sources: For instance, composition measurements may be provided online by several gas chromatographs, in addition to laboratory analysis performed at irregular and infrequent intervals. Viscosity can be measured online via viscometers; however, a faster and more-robust indication may be available through measuring the pressure drop across a pipe. Both online measurements are typically augmented with infrequent laboratory analyses.

Recently, there has been growing interest in the use of multiple sensors to increase the capabilities of intelligent systems,[1] as well as to improve the robustness of complex dynamical systems. Typical domains of application for sensor fusion techniques are robotics[2] and inertial navigation,[3] as well as target detection and tracking.[4] Applications in the process industries typically involve the related problems of inferential measurements,[5,6] data reconciliation,[7,8] and sensor validation,[9] but rarely with the sensor fusion problem stated previously.

Generally, sensor fusion may be considered as a methodology for integrating signals from multiple sources,[10] with the following potential advantages: information from redundant or complementary sensors

* To whom correspondence is to be addressed. Tel.: (302)-831-4504. Fax: (302)-831-1048. E-mail: ogunnaike@che.Udel.edu.

10.1021/ie049614t CCC: $30.25
Published on Web 12/09/2004

can be utilized, resulting in reduced uncertainty, and increased precision and reliability in the event of a sensor failure. Moreover, information may be provided with a smaller delay or at a higher rate by multiple sensor arrangements (e.g., multiple online analyzers) than by a single sensor. The common principle utilized by most sensor fusion techniques is that of measurement redundancy, which can be present in two forms:

(1) Direct Redundancy. Multiple instruments provide estimates of the same process variable. However, these redundant measurements may be available at different frequencies or with different time delays. This type of redundancy is the focus of this article.

(2) Analytical Redundancy. In addition to a direct measurement, redundant information about a process variable may be available from measurements of other "analytically" related variables.

Sensor fusion methods range from applications of probabilistic models (such as Bayesian reasoning), least-squares techniques (such as Kalman filtering), to artifical intelligence approaches.[10] In what follows, we briefly review several approaches that are currently available for sensor fusion and fault detection in directly redundant sensor systems.

If more than two redundant measurements of a process variable are available, simple "voting logic" can be used to eliminate faulty sensors. In this approach, sensors that do not agree with the majority of the remaining sensors, are considered faulty (majority voting). However, voting sytems perform poorly in detecting "soft failures", such as small sensor biases.[11]

As an extension to voting logic, fuzzy logic techniques can be applied to data fusion problems. Fuzzy logic involves extensions of Boolean logic to a continuous-valued logic via the concept of membership functions.[12] For instance, Hong and Wang[13] proposed a technique for fusing both noisy and fuzzy data, which uses both Kalman filtering and fuzzy arithmetic.

Expert system approaches have been proposed in the context of sensor validation. The approach of Lee[14] is based on causal relations between sensors. In particular, the expected value of a sensor output can be determined from measurements of all the remaining sensors when they are in a causal relation. Individual sensors are assigned validity levels corresponding to the number of independent sources of supportive evidence that a particular sensor reading is valid.[14]

Applications of artificial neural networks (ANNs) in sensor fusion and failure detection are also quite common.[6,15] For instance, ANNs have been used[6] to combine measurements from several analyzers and standard reactor measurements to predict key process variables for online monitoring of a fermentation process. Naidu et al.[15] used a back-propagation network for sensor failure detection based on Fourier analysis of a diagnostic signal.

Kalman filtering provides a convenient framework for performing sensor fusion, and several applications have been reported in the literature.[5,16,17] For instance, a Kalman filter has been used in the context of inferential estimation of a process variable from secondary measurements that are available more frequently than the primary measurement.[5] In another approach, a main Kalman filter that processes all measurements is used in conjunction with a bank of auxiliary Kalman filters that process subsets of the measurements;[16] failure detection is performed by checking the consistency between the state estimates of the main Kalman filter and the corresponding estimates of the auxiliary filters. This technique is based on the premise that sensor failures have only a small effect on the overall estimate provided by the main Kalman filter, whereas the effects on the auxiliary filters that contain the faulty sensors will be more significant. Similarly, decentralized Kalman filters have been shown to have improved fault tolerance, when compared to a single ("centralized") Kalman filter.[17]

Yet to be addressed in a coherent, systematic manner is the problem of combining measurements from different sources in an actual industrial setting, where sensor characteristics are not known perfectly, and robustness to sensor biases and other partial sensor failures are critical to process operation. This is the focus of this paper. First, we address the nominal sensor fusion problem and then explicitly address the robust problem by augmenting the nominal approach with sensor failure detection, classification, and compensation schemes based on statistical hypothesis testing.

The remainder of this article is organized as follows. In section 2, we discuss an example of sensor fusion in biological systems and identify one of the underlying principles that may be utilized for process applications. Section 3 presents a dynamic sensor fusion approach based on Kalman filtering, whereas section 4 examines robustness issues in the presence of sensor failures. The application study presented in section 5 demonstrates the performance of the proposed techniques on an industrial copolymerization reactor.

2. Sensor Fusion in Biological Systems

Biological control systems face problems that are similar to those encountered in chemical process control. However, biological systems seem to solve these problems much more efficiently than currently available industrial control approaches. Current knowledge of biological control systems indicates that the high level of reliability and accuracy required of measurements of critical physiological variables is achieved by highly integrated measurement schemes with large arrays of partially redundant sensors. Moreover, these systems are highly adaptive and remarkably tolerant to sensor and actuator failure. Thus, biological control systems may provide a viable source of inspiration for developing sensor fusion techniques. Referring to the fault-tolerant signal processing capabilities of neural systems, it has been recognized[18] that "*robust neural mechanisms employ a combination of (i) a large number of inputs carrying overlapping but nonidentical information, and (ii) a limitation of the effect that each input can have on the output by a suitable nonlinearity.*"

In the following, we will discuss an example of sensor fusion in biological control systems and note several of its characteristics that may be responsible for achieving properties (i) and (ii) cited in the previous paragraph.

2.1. The Baroreceptor Reflex. The baroreceptor reflex (or baroreflex for short) is responsible for short-term blood pressure regulation. In the baroreflex, arrays of pressure sensitive neuronal stretch receptor endings located in the aortic arch and carotid sinus transmit blood pressure information to other neurons (so-called "second-order neurons") in the brain. Information is transmitted via voltage signals that have the form of

trains of voltage spikes of very short duration, which are referenced as "action potentials". Whenever such a voltage spike is generated, a neuron is said to fire an action potential. The frequency with which these action potentials are generated is directly related to the pressure at the location of the receptor: If the blood pressure is low, the firing frequency of these baroreceptive neurons will be low, whereas for higher pressures, the firing frequency will increase, up to a maximum frequency (or saturation frequency).

In a particular region in the brain—the so-called NTS (nucleus tractus solitarius)—these baroreceptor signals are integrated with other sensory signals that indicate blood composition and other physiological variables. These integrated signals are then processed further by two parallel control systems that eventually send signals to the heart, to maintain the blood pressure at a desired level. See refs 19 and 20 for a more-detailed discussion of the baroreceptor reflex and potential applications in chemical process control.

Several properties of the baroreceptor reflex seem to be particularly relevant in the context of sensor fusion:

(1) The baroreceptive neurons are organized in the form of a large number of neurons with narrow individual operating ranges, which have the highest sensitivity around a nominal operating point (i.e., pressure) and decreasing sensitivity away from this nominal operating point. The individual baroreceptors have partially overlapping operating ranges such that the desired measurement range is completely spanned by the receptor ensemble rather than by single neurons.

(2) The frequency with which a baroreceptor can generate and transmit action potentials is bounded between a lower limit (the frequency at which the receptor generates purely random action potentials that are unrelated to the blood pressure) and an upper limit at which the receptor saturates. Beyond this saturation limit, an increase in pressure does not further elevate the firing frequency, resulting in a sigmoidal-shaped firing characteristic as shown in ref 21.

(3) The baroreflex possesses a high degree of robustness, because the removal of a significant fraction of baroreceptive neurons has a relatively small effect on the overall control system performance.[22]

(4) The baroreceptor reflex apparently performs outlier detection; data that are inconsistent with other sensory inputs are "ignored" at the central signal processing stage.[22]

Note that property 1 provides the brain with a large number of inputs carrying overlapping but nonidentical information about the blood pressure. The steady-state firing characteristics of the individual receptors, which essentially have a sigmoidal shape, as discussed in the previously mentioned property 2, can be thought of as representing a nonlinearity that can limit the effect that each input (a baroreceptor signal) can have on the output (the next signal processing stage).

Unlike the baroreceptors, whose pressure−response functions have been well-defined, very little is known about the processing of the baroreceptor signals within the brain. The first stage of this signal processing occurs in the so-called second-order neurons in the NTS. Electrophysiological recordings were undertaken by Robert F. Rogers at DuPont to define the properties of these neurons. The results of this work indicate that second-order neurons receiving baroreceptor inputs are capable of providing the direction and magnitude of pressure changes but are unlikely to provide a reliable encoding of mean pressure.[23,24] These findings result in the hypothesis that the central nervous system extracts and separates different pieces of information provided within it and uses them for different functions.[24]

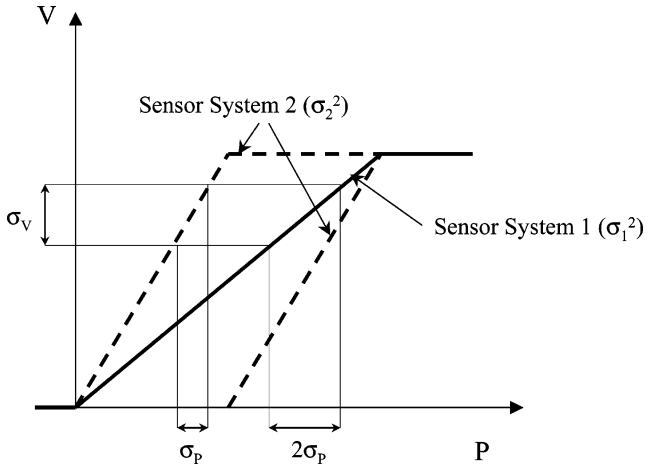

Figure 1. Multisensor systems.

Despite these recent findings, the actual mechanisms for signal processing within this part of the brain are still poorly understood. Therefore, our objective is not to "copy" the neuronal signal processing within the NTS, but rather to mimic some of the strategies or general principles (e.g., architecture) of the baroreceptor reflex. As an example of this approach, we discuss in the following paragraphs how one of the general characteristics of the baroreceptor reflex (the utilization of large numbers of sensors with narrow individual operating ranges) can be exploited in designing multisensor systems.

2.2. Idealized Sensor Configurations. Here, we consider a multisensor system, which can be thought of as being derived from the multisensor array seen in the baroreflex by making the following crude approximations: (i) the sigmoidal sensor characteristics are replaced by piecewise linear ones, and (ii) there is no overlap between the individual sensor operating ranges (i.e., the entire operating range is divided into segments of identical width and every sensor provides measurements only in its assigned local operating range). To quantify the advantages of such a multisensor system over a sensor system in which all sensors cover the entire measurement range, we consider the following example. We compare two sensor systems that each consist of two sensors (see Figure 1). Sensor system 1 utilizes two sensors that have the same global operating range; its sensors are called Type 1 sensors. Sensor system 2 consists of two sensors, one of which covers the lower half of the overall operating range and the other sensor covers the upper half. Its sensors are called Type 2 sensors.

Figure 1 shows the raw sensor output (for example, a voltage), as a function of the process variable P, for both the sensors of sensor system 1 (solid line) and sensor system 2 (dashed lines). If we assume that, independent of the sensor type, the same raw measurement error is present (characterized by the standard deviation σ_V), the standard deviation of the estimated process variable that results from a Type 2 sensor (i.e., σ_P) is only half the standard deviation associated with a Type 1 sensor ($2\sigma_P$). The variance of the composite measurement of sensor system 1 (σ_1^2) (obtained as the

average of the two Type 1 measurements) is given by

$$\sigma_1^2 = 2\sigma_P^2 \tag{1}$$

whereas the variance σ_2^2 of the "combined" sensor system 2 measurement (which is simply obtained from one of the two Type 2 sensors) is

$$\sigma_2^2 = \sigma_P^2 \tag{2}$$

Thus, the advantage of limited operating ranges is better precision. However, note that, in this example, there is no overlap between the sensor operating ranges. Therefore, the reliability of sensor system 1 can be expected to be superior to that of sensor system 2, because there are two sensors available at every point of the input range. The advantage of overlapping sensor operating ranges is that redundancy is introduced and, therefore, the robustness of the measurement system is improved. This leads to the conclusion that, with a combination of partially overlapping sensors with limited operating ranges, we can either increase robustness at the expense of precision or increase precision at the expense of robustness.

These concepts obviously have significant implications for sensor design (choosing the number and configuration of sensors required to achieve specified precision and robustness goals). However, they also result in the complementary "analysis problem": how does one systematically combine the multiple measurements from a given multisensor configuration? How do the various design variables affect achievable precision and robustness? A detailed discussion of these and other issues concerning the analysis of multisensor configurations is provided in the next section.

3. A Dynamic Sensor Fusion Approach

The sensor fusion approach developed in this section requires two types of information: (i) a *process model*, i.e., a description of how the true process variable η evolves over time, and (ii) *sensor models*, i.e., representations of how the individual sensors respond to changes in the process variable η.

3.1. Process Model. The process model considered here is the general linear state-space model of the form

$$x_p(k+1) = A_p(k)x_p(k) + B_p(k)u(k) + F_p(k)v_p(k) \tag{3}$$

$$\eta(k) = C_p(k)x_p(k) \tag{4}$$

where x_p is the state vector of the process model, u a vector of process inputs, v_p an uncorrelated, white (i.e., independent, identically distributed) zero-mean random disturbance, and η the true value of the process variable. Output disturbances are not included, because η represents the true (noise-free) value of the process variable. We will subsequently use the notion of a "precise process model" if the effects of the stochastic model components on the process variable η are small, whereas a model will be referenced as being "imprecise" if the variance of the stochastic components at the output is large.

In cases where a complete process model that contains deterministic and stochastic components as in eqs 3 and 4 is not available, the process variable η can be described as a purely stochastic process. A generally accepted and nonrestrictive assumption is that $\eta(k)$ varies relatively smoothly with k. A natural class of such smoothly varying random processes is the class of *martingales*,[25] for which the following is true:

$$E[\eta(k) \mid \eta(k-1), ..., \eta(1)] = \eta(k-1) \tag{5}$$

The simplest martingale is the random walk process, which is defined by

$$\eta(k+1) = \eta(k) + v_p(k) \tag{6}$$

where $v_p(k)$ is an uncorrelated, white, zero-mean random sequence with a variance of $E[v_p(k)^2] = P_p(k)$. Note that this process model has been introduced here as one that, in the absence of a deterministic model, requires the least amount of extraneous detail to describe a smoothly varying random process. Moreover, because a random walk process is nonstationary, the estimate of η that is based on this model and derived by the optimal filter discussed below will be unbiased, if the actual process variable is asymptotically constant.

In practice, some knowledge about the process time constants and operating conditions can provide guidance in selecting adequate values for $P_p(k)$. By choosing a very small value for P_p, the estimate will rely heavily on the smooth variation model implied by eq 6, whereas, by choosing very large values for P_p (relative to the precision of the sensors), the estimate of η will mainly rely on measurements.

3.2. Sensor Models. The individual sensors are assumed to be linear and can be described by standard linear state-space models. Examples of such sensor models include the following:

(a) Unbiased Static Sensor. The measurement uncertainty is completely captured by additive measurement noise, and the sensor model consists of a single output equation:

$$y_i(k) = \eta(k) + w_i(k) \tag{7}$$

(b) Static Sensor with Constant Bias. A bias state is introduced as

$$\beta_i(k+1) = \beta_i(k) \tag{8}$$

and the output (as defined by eq 7) is modified to include the bias state,

$$y_i(k) = \eta(k) + \beta_i(k) + w_i(k) \tag{9}$$

(c) Static Sensor with Time-Varying Bias. The simplest representation of a time-varying bias is again the random walk process, which leads to the following sensor model:

$$\beta_i(k+1) = \beta_i(k) + v_i(k) \tag{10}$$

$$y_i(k) = \eta(k) + \beta_i(k) + w_i(k) \tag{11}$$

where $E[v_i(k)^2]$ characterizes the "severity" of the random walk.

Next, we combine the individual sensor models into a single model of the corresponding multisensor system by defining the state vector,

$$x_s(k) = [x_1(k)^T \ \cdots \ x_n(k)^T]^T \tag{12}$$

Explicitly indicating the potential time dependence of the system matrixes, the model of the multisensor system can be written as

$$x_s(k+1) = A_s(k)x_s(k) + B_s(k)\eta(k) + F_s(k)v_s(k) \quad (13)$$

$$y(k) = C_s(k)x_s(k) + D_s(k)\eta(k) + w(k) \quad (14)$$

where v_s and w are uncorrelated, white, zero-mean random disturbances.

3.3. Solution of the Dynamic Sensor Fusion Problem. The process and sensor models can be combined by introducing the following matrices:

$$A(k) = \begin{bmatrix} A_s(k) & B_s(k)C_p(k) \\ 0 & A_p(k) \end{bmatrix}$$

$$B(k) = \begin{bmatrix} 0 \\ B_p(k) \end{bmatrix}$$

$$C(k) = [C_s(k) \;\; D_s(k)C_p(k)] \quad (15)$$

It is entirely reasonable to assume that the process and sensor disturbances are uncorrelated. Hence,

$$F(k) = \begin{bmatrix} F_s(k) & 0 \\ 0 & F_p(k) \end{bmatrix} \quad (16)$$

with the overall state and disturbance vectors defined as

$$x(k) = \begin{bmatrix} x_s(k) \\ x_p(k) \end{bmatrix}$$

$$v(k) = \begin{bmatrix} v_s(k) \\ v_p(k) \end{bmatrix} \quad (17)$$

the combined model can be written as

$$x(k+1) = A(k)x(k) + B(k)u(k) + F(k)v(k) \quad (18)$$

$$y(k) = C(k)x(k) + w(k) \quad (19)$$

with

$$\mathrm{cov}[v(k), v(k)] = P(k)$$

$$\mathrm{cov}[w(k), w(k)] = Q(k)$$

$$x(0) = x_0 \quad (20)$$

We also assume that the measurement noise ($w(k)$) and the state disturbances ($v(k)$) are uncorrelated, i.e., cov$[v(k), w(k)] = 0$. Based on the description of the overall system, the Kalman filter[26] provides the optimal state estimate (i.e., the estimate with the minimum estimation error variance among all unbiased estimates). It can be written in the following form:

$$\hat{x}(k+1|k+1) = A(k)\hat{x}(k|k) + B(k)u(k) + K(k+1) \times [y(k+1) - C(k+1)A(k)\hat{x}(k|k)] \quad (21)$$

$$\Sigma(k+1) = A(k)\{\Sigma(k) - \Sigma(k)C(k)^{\mathrm{T}}[C(k)\Sigma(k)C(k)^{\mathrm{T}} + Q(k)]^{-1}C(k)\Sigma(k)\}A(k)^{\mathrm{T}} + F(k)P(k)F(k)^{\mathrm{T}} \quad (22)$$

$$K(k+1) = \Sigma(k+1)C(k+1)^{\mathrm{T}}[C(k+1)\Sigma(k+1)C(k+1)^{\mathrm{T}} + Q(k+1)]^{-1} \quad (23)$$

where K is the Kalman filter gain and Σ (with $\Sigma(0) = \Sigma_0$) is the so-called a priori covariance of the state estimation error; it is based on measurements up to $y(k)$ but not including $y(k+1)$. Equivalent Kalman filter formulations that are based on an a posteriori covariance, which includes knowledge of the most-recent measurements (i.e., up to $y(k+1)$), are also available. The a posteriori estimation error covariance T is related to the a priori covariance, via

$$T(k) = [I - K(k)C(k)]\Sigma(k) \quad (24)$$

Recall that the true process variable $\eta(k)$ is a known linear combination of a subset of the state variables in eq 17, as stated in eq 4. Therefore, the optimal estimate of η is obtained as a linear combination of the state estimates provided by the Kalman filter.

3.4. Practical Considerations. We now consider several specific practical situations that the sensor fusion approach will face. In particular, the following cases will be addressed:

(a) Static Sensors. The sensor dynamics can be neglected, resulting in a sensor model as described by eq 7.

(b) Delayed Sensors. The sensors provide delayed measurements but are otherwise static.

(c) Multirate Sensors. In many applications, some of the available measurement devices (e.g., a chemical analyzer) do not provide measurements at every sampling instant but at multiples of the basic sampling interval. Although such instruments also often produce delayed measurements, we will only consider undelayed multirate sensors, for the sake of clarity.

(d) Reference Sensor. In addition to online sensors, a laboratory analysis of a process sample is performed at infrequent and irregular time intervals, which provides an unbiased "reference" measurement.

3.4.1. Static Sensors. If we assume that the true process variable η is generated by a random walk (see eq 6), the combined system consisting of process and sensor models (eq 7) can be written as

$$\eta(k+1) = \eta(k) + v(k) \quad (25)$$

$$y(k) = C\eta(k) + w(k) = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}\eta(k) + w(k) \quad (26)$$

with

$$\mathrm{cov}[v(k), v(k)] = P$$

$$\mathrm{cov}[w(k), w(k)] = Q = \begin{bmatrix} Q_1 & & \\ & \ddots & \\ & & Q_n \end{bmatrix} \quad (27)$$

Note that if there was a deterministic model component, it could be omitted entirely for investigations concerning estimation precision.[27] Therefore, the results derived here are also valid if there are underlying deterministic models, whose uncertainties can be described by eq 6.

The Kalman filter equations (eqs 21−23) can now be used to determine the estimate of the process variable, $\hat{\eta}(k)$. Because the multisensor system in eqs 25−27 is time-invariant, the Kalman filter gain and covariance will converge to their steady-state values, under reasonably mild conditions (for details, see ref 28, pp 252 ff). In this situation, it is common practice to use the steady-state Kalman filter for state estimation, obtained by iterating eq 22, or directly by letting $\Sigma(k+1) = \Sigma(k) =$

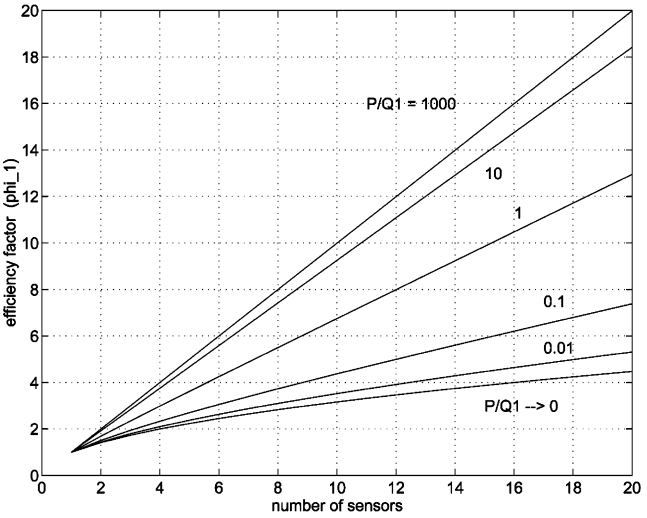

Figure 2. Effect of adding redundant sensors.

Σ in eq 22, which results in the algebraic equation (with Σ being scalar)

$$\Sigma C^{\mathrm{T}}(C\Sigma C^{\mathrm{T}} + Q)^{-1}C\Sigma = P \tag{28}$$

The solution of eq 28 immediately yields the steady-state Kalman filter gain K, as

$$K = \Sigma C^{\mathrm{T}}[C\Sigma C^{\mathrm{T}} + Q]^{-1} \tag{29}$$

The steady-state a posteriori estimation error variance T is given by

$$T = (1 - KC)\Sigma \tag{30}$$

These results are now used to determine the performance gain that can be achieved using redundant sensors with identical characteristics. For this purpose, we define an efficiency factor ϕ_1:

$$\phi_1 = \frac{{\sigma_1}^2}{{\sigma_n}^2} = \frac{T(1\ \text{sensor})}{T(n\ \text{sensors})} \tag{31}$$

where ${\sigma_i}^2$ is the error variance associated with estimates provided by an ensemble of i identical sensors. Figure 2 shows the efficiency factor ϕ_1, as a function of the total number of sensors. Not surprisingly, the potential improvement is dependent on the ratio P/Q_1, which characterizes the precision of the sensors, relative to the precision of the process model. In particular, if there is no process knowledge available (i.e., $P/Q_1 \rightarrow \infty$), the efficiency factor increases linearly with the total number of sensors. Because, in this case, the optimal estimate is simply obtained by averaging the individual measurements, it follows that

$$\phi_1 = n \tag{32}$$

As the process model becomes more precise, compared to the sensors, the advantages of using multiple sensors diminish gradually. For the limiting case $P/Q_1 \rightarrow 0$, the efficiency factor ϕ_1 reduces to

$$\phi_1 = \sqrt{n} \tag{33}$$

3.4.2. Delayed Sensors. Here, we consider delayed but otherwise static sensors. Starting with a random walk process model (eq 25) with $E[v_1(k)^2] = P_1$, and one undelayed static sensor as in eq 7, with $E[w_1(k)^2] = Q_1$, we now add a second static sensor with a measurement delay of θ sampling periods:

$$y_2(k) = \eta(k - \theta) + w_2(k) \tag{34}$$

with $E[w_2(k)^2] = Q_2$. Introducing the following state vector,

$$x(k) = [\eta(k)\ \ \eta(k-1)\ \ \cdots\ \ \eta(k-\theta)]^{\mathrm{T}} \tag{35}$$

yields the overall system description

$$x(k+1) = \begin{bmatrix} 1 & & & \\ 1 & \ddots & & \\ & \ddots & & \\ & & 1 & 0 \end{bmatrix} x(k) + \begin{bmatrix} v_1(k) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{36}$$

$$y(k) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix} x(k) + w(k) \tag{37}$$

where the noise terms are characterized by

$$\mathrm{cov}[v(k), v(k)] = P = \begin{bmatrix} P_1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}$$

$$\mathrm{cov}[w(k), w(k)] = Q = \begin{bmatrix} Q_1 & \\ & Q_2 \end{bmatrix} \tag{38}$$

The steady-state Kalman filter equations are again used to assess the improvement in the estimation precision that can be achieved by adding a delayed sensor. For this purpose, an efficiency measure (ϕ_2) is introduced:

$$\phi_2 = \frac{{\sigma_1}^2}{{\sigma_\theta}^2} \tag{39}$$

where ${\sigma_1}^2$ is the error variance in estimating η if a single undelayed sensor is used and ${\sigma_\theta}^2$ is the corresponding variance if a second sensor with a delay θ is added.

The results for the case, $Q_1/Q_2 = 10$ are shown in Figure 3. The efficiency factor ϕ_2 is shown as a function of the time delay of the secondary sensor, for five different values of the P_1/Q_1 ratio, which characterizes the relative precision of the primary sensor versus the precision of the process model. The results show that adding a delayed sensor is advantageous only if there is some process knowledge available, because a delayed sensor can only contribute to the estimate of the process variable if reliable prediction beyond the delay is possible. Moreover, the improvements that can be obtained from adding a delayed sensor increase with increasing precision of the process model (i.e., decreasing P_1/Q_1 ratio).

3.4.3. Multirate Sensors. In addition to a primary sensor, which provides measurements at every sampling instant, one multirate sensor is available that provides measurements every M sampling instants, where M is the period of the multirate sensor. For a discussion of multirate state estimation, see, e.g., ref 29.

Here, again, we assume a stochastic process model of the form of eq 25. Because the sensors are assumed

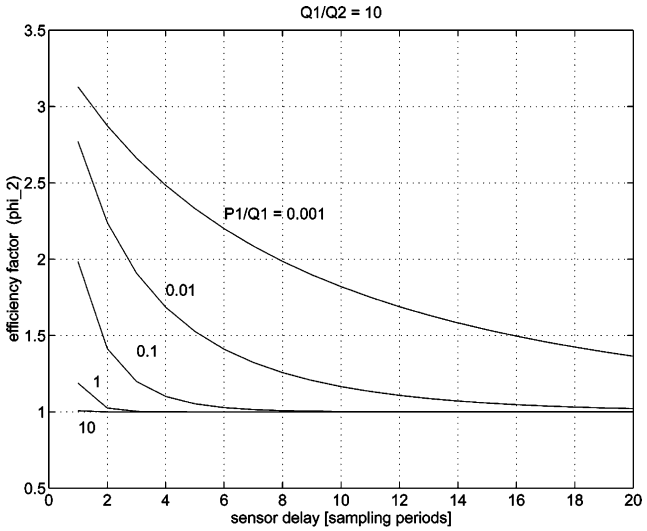

Figure 3. Effect of adding a delayed sensor ($Q_1/Q_2 = 10$).

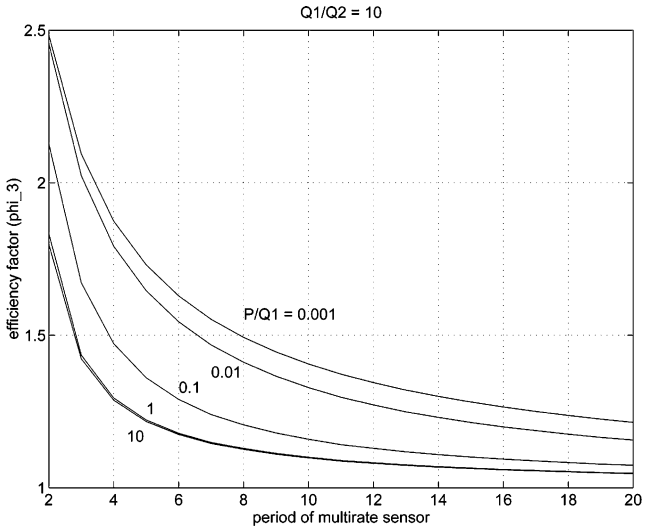

Figure 4. Effect of adding a multirate sensor ($Q_1/Q_2 = 10$).

to be static (and do not exhibit any time delays), η is the only state variable of the combined process and sensor model. The output equation is dependent on whether one or both measurements are available at a particular sampling instant. Thus, we obtain the following time-varying system description:

$$\eta(k+1) = \eta(k) + v(k) \tag{40}$$

$$y(k) = C(k)\eta(k) + w(k) \tag{41}$$

with

$$\mathrm{cov}[v(k), v(k)] = P$$

$$\mathrm{cov}[w(k), w(k)] = Q = \begin{bmatrix} Q_1 & \\ & Q_2 \end{bmatrix} \tag{42}$$

where $y(k) = [y_1(k)\ y_2(k)]^T$, $w(k) = [w_1(k)\ w_2(k)]^T$, and $C(k)$ is a periodically time-varying matrix,

$$C(k) = \begin{cases} [1\ \ 1]^T & (\text{for } k = nM,\ n \in \mathcal{N}) \\ [1\ \ 0]^T & (\text{otherwise}) \end{cases} \tag{43}$$

Because of the periodicity of $C(k)$, the Kalman filter converges to a stable periodic filter after the initial transients have died out. It is this stable periodic solution that can again be used to assess the performance of the multisensor system.

The periodic $\Sigma(k)$ is obtained by iterating eq 22 until it converges to a stable periodic solution; and the periodic solutions for $K(k)$, and $T(k)$, then follow directly from eqs 23 and 24. For periodic Kalman filter applications, also see refs 30 and 31. Alternatively, the periodic time-varying system could be "lifted" into a time-invariant system by considering the output as consisting of all measurements made during one sensor period.[27] This latter approach provides computational advantages, because the steady-state Kalman filter is determined directly for the lifted time-invariant system, instead of iteratively from eq 22. However, because computational aspects are of minor concern for this (offline) study, we adopt the approach of iterating eq 22.

Again, an efficiency factor (ϕ_3) is introduced, as

$$\phi_3 = \frac{{\sigma_1}^2}{{\bar{\sigma}_M}^2} \tag{44}$$

where ${\sigma_1}^2$ is the estimation error variance if a single sensor with sensor period $M = 1$ is used, and ${\bar{\sigma}_M}^2$ is the average value of the variance if a multirate sensor with period $M > 1$ is added to the primary sensor. The average ${\bar{\sigma}_M}^2$ has been introduced, because the multisensor system is now periodic and obtaining a meaningful performance measure requires an averaging of the estimation error variance over one period M of the multirate sensor after eq 22 has converged.

Figure 4 shows, for $Q_1/Q_2 = 10$, the efficiency factor ϕ_3, as a function of the period of the multirate sensor, for different values of P/Q_1. It can be concluded that the addition of a multirate sensor is always advantageous. Even if there is no process knowledge available ($P/Q_1 \rightarrow \infty$), the additional measurements provided by the multirate sensor every M sampling instants will reduce the estimation error variance. Naturally, these improvements diminish with increasing sampling period of the multirate sensor. In analogy to the case of adding a delayed sensor, the improvements are dependent on the precision of the process model, because of the need for predicting the process variables during the intersample period between multirate measurements.

3.4.4. Reference Sensors. The key characteristics of a reference sensor are (i) it is static and unbiased, (ii) its time delay is large, compared to the sampling interval of the online sensors, and (iii) the measurement time delay is not necessarily constant, because the time needed to analyze a process sample may vary. Given property (ii), one could consider adding as many delayed versions of $\eta(k)$ as state variables as called for by the maximum possible time delay associated with the reference sensor. However, this would be impractical, because an excessive numbers of state variables may be required in some cases. Instead, we propose to incorporate the reference sensor in the following manner:

(1) The original system description (without the reference sensor) is augmented by the output equation corresponding to the reference sensor,

$$y_{\mathrm{ref}} = \eta(k) + w_{\mathrm{ref}}(k) \tag{45}$$

and the measurement noise covariance matrix Q is augmented by one additional diagonal element, $Q_{\mathrm{ref}} = E[w_{\mathrm{ref}}(k)^2]$. A time-varying output matrix $C(k)$ is used to indicate when a reference measurement is available.

(2) At time instants at which a sample for a laboratory measurement is taken (for example, k_{ref}), the following variables are stored in memory:

$$k_{\mathrm{ref}}, y(k_{\mathrm{ref}}), C(k_{\mathrm{ref}} - 1), \Sigma(k_{\mathrm{ref}}), \hat{x}(k_{\mathrm{ref}} - 1|\, k_{\mathrm{ref}} - 1) \tag{46}$$

In addition, all measurements that are obtained from instant k_{ref} up until the reference measurement becomes available, have to be stored in memory.

(3) At sampling instants at which a reference measurement becomes available (for example, k), we "rewind" the Kalman filter to sampling instant k_{ref} by retrieving the aforementioned variables from memory. A new output matrix $C(k_{\mathrm{ref}})$ that takes into account the now-available reference measurement is obtained, and the filter gain $K(k_{\mathrm{ref}})$ and state estimates $\hat{x}(k_{\mathrm{ref}}|k_{\mathrm{ref}})$ are recalculated.

(4) Based on the updated Kalman filter at time k_{ref}, and all measurements obtained between k_{ref} and k, the filter is then iterated forward to the current sampling instant k.

This procedure generates a new sequence of state estimates, including estimates of the process variable, for the time period between taking the sample and the current sampling instant. Because reference measurements are available rather infrequently, this approach can be expected to be computationally more efficient than the approach of introducing a potentially large number of additional state variables.

4. Robustness Issues

The Kalman filter is based on assumptions (e.g., uncorrelated random noise) that, quite often, are violated in practice. Therefore, careful attention must be given to issues concerning robustness in the presence of any deviations from nominal system behavior. In particular, the following situations will arise frequently:

(a) Measurement Outliers. Measurements with "unusually large" errors constitute deviations from the nominal sensor description.

(b) Sensor Degradation. The performance of individual sensors may degrade over time, i.e., the measurement error variances may increase and/or the bias characteristics may change considerably.

(c) Changes in the Process Behavior. The filter is based on a particular process model. However, the process may change its dynamic behavior over time, as periods of operation close to a particular operating point may be followed by periods of frequent process changes and transitions.

Robustness in the face of the aforementioned scenarios is achieved via failure detection, classification, and compensation techniques, as discussed below.

4.1. Failure Detection. The innovations $\omega(k)$ are defined as the difference between sensor readings at time instant k and the predicted values based on information up to $k - 1$:

$$\begin{aligned}\omega(k) &= y(k) - C(k)\hat{x}(k|k-1) \\ &= y(k) - C(k)A(k)\hat{x}(k-1|k-1)\end{aligned} \tag{47}$$

If there is no mismatch between model assumptions and the true system, the innovation sequence is a white noise sequence with known covariance. However, if the true process does not match the process model, then the Kalman filter loses its optimality and the innovations will be correlated. Therefore, the innovations can be used for failure detection or for simply detecting changes in the process; for instance, tests for whiteness of the innovations, or tests that detect shifts in the mean or increases in the variance of the innovations, have been suggested.[32]

Here, we consider the so-called standardized innovations,

$$\tilde{\omega}(k) = [C(k)\Sigma(k)C(k)^{\mathrm{T}} + Q]^{-1/2}\omega(k) \tag{48}$$

which provide a normalization such that $\mathrm{cov}[\tilde{\omega}(k), \tilde{\omega}(k)] = I$, where I denotes the identity matrix.

The standardized innovations are more convenient for use in fault detection, because the statistical properties of the standardized innovations remain constant, even though the system description may be time-varying. Any deviations of the true system behavior from the model description will be reflected in changes in the statistical properties of the (standardized) innovations. If such changes are detected, it can be concluded that a failure has occurred somewhere in the system. However, the more specific task of isolating the failure origin from the innovations is a rather difficult problem in general. Yet, in the present application, because of the particular structure of multisensor systems, which can be characterized by exhibiting very weak interactions between the model components associated with the individual sensors, we will make the assumption that a change in the statistical properties in one component of the innovation sequence can be associated with a fault in the corresponding sensor. Because any changes in the process behavior that are not captured by the process model would also affect the innovations, we must monitor the adequacy of the (stochastic) process model. In particular, we have to detect rapid process changes, because such changes would affect the statistical properties of the innovations, and false indications of sensor failures could occur.

4.1.1. Process Model Validity. If a purely stochastic process model is used, the problem of monitoring the model validity involves estimating the covariance matrix that is associated with the process model P_{p}. For a random walk process, which is a white noise process, the following approach can be used for estimating P_{p}: The model described by eq 6 is completely characterized by a single parameter, $P_{\mathrm{p}} = E[v_{\mathrm{p}}(k)^2]$, which can be estimated online. For this model, we obtain

$$\begin{aligned}E[(\eta(k) - \eta(k-r))^2] = E[(v_{\mathrm{p}}(k) + v_{\mathrm{p}}(k-1) + \cdots + \\ v_{\mathrm{p}}(k-r))^2] = rP_{\mathrm{p}}\end{aligned} \tag{49}$$

Assuming for the moment that the process model is adequate, the combined measurement obtained from the Kalman filter is an unbiased estimate of the true process

variable. Therefore, we replace $\eta(k - i)$ with the corresponding estimates to obtain

$$E[(\hat{\eta}(k) - \hat{\eta}(k - r))^2] = rP_p + 2\sigma_\eta^2 \tag{50}$$

where σ_η^2 is the estimation error variance. Note that σ_η can be time-varying (due to the possible time-varying nature of the system description) and may not be perfectly known. Fortunately, its value is not required for estimating P_p, if a sufficiently large value for r is used. For large values of r, rP_p dominates over $2\sigma_\eta^2$ and eq 50 can be used to derive an upper bound for P_p from a sequence of η estimates of length N, according to

$$\hat{P}_p(k) \approx \frac{\sum_{k=r+1}^{N} [\hat{\eta}(k) - \hat{\eta}(k - r)]^2}{(N - r)r} \tag{51}$$

4.1.2. Detection Algorithm. To detect sensor failures, and, at the same time, to be able to track rapid process changes, the following tests are implemented.

(1) Outlier Test. Whenever the standardized innovation associated with a particular sensor is "unusually large", the corresponding measurement is very likely to be an outlier. This leads to the following test: Sensor i is considered faulty if

$$||\tilde{\omega}_i(k)|| \geq \alpha_i \tag{52}$$

where α_i is an appropriately chosen threshold. Knowing that, under the assumption of a perfect model, the variances of the standardized innovation sequences are unity, the thresholds α_i can be easily chosen a priori.

(2) Bias and Variance Tests. Biases whose magnitudes are smaller than the thresholds for outlier detection can be detected by analyzing sequences of past innovations (with any outliers detected in test 1 removed from the sequence). The statistical test used to detect shifts in the mean of a random variable is described in Appendix A.1. Note that this test assumes that the variance of the sequence is known—which is the case here, at least under the perfect model assumptions. However, because there will always be some process/model mismatch, the assumed variances (denoted as $\{\gamma_i\}$ for future reference) should rather be viewed as tuning parameters for the sensor fusion approach. They can be determined by observing the innovations of the Kalman filter for actual process data, i.e., by processing a "test" data set.

The innovation sequence can also be used to detect sensor degradation, because an increase in the measurement error variance of one sensor will primarily result in an increase in the variance of the corresponding innovation sequence $\tilde{\omega}_i$. To detect this type of sensor degradation, we use the test described in Appendix A.2, which again requires the specification of the assumed variances of the standardized innovations.

(3) Model Validity Test. The outcome of the aforementioned tests is dependent on the validity of the model characterizing the dynamics of the true process variable. For the case of the random walk process model, we monitor the model validity by estimating P_p online, using eq 51. The process model is considered inadequate if

$$\hat{P}_p(k) > P_p^0 \tag{53}$$

where P_p^0 is the nominal value of P_p used by the Kalman filter. Note that we only consider the model inadequate if the estimated variance is larger than that assumed by the model. Although the Kalman filter will also be sub-optimal if the actual process variance P_p is smaller than P_p^0, we consider P_p^0 as a lower bound, to ensure that the filter remains "alert" and can respond quickly to process changes, even though the process variable might have hardly changed for an extended period of time.

If the process model is determined to be inadequate, the system description is updated by introducing the estimate $\hat{P}_p(k)$, instead of P_p^0, in matrix P in eq 20. Furthermore, the detection mechanisms for soft sensor failures (test 2 previously described) are disabled temporarily, because inadequate model assumptions are very likely to affect the innovations. Note that, although the process description is updated whenever a rapid process change is detected, the process model may still be inadequate, because sudden changes in the process dynamics cannot be captured immediately by $\hat{P}_p$. However, to be robust in the face of gross measurement outliers, the outlier test (test 1) is still performed.

(4) Test on Bias Estimates. As an additional robustness measure, we also monitor the bias estimates that are provided by the Kalman filter. A very simple failure detection scheme is to consider a sensor to be faulty if its bias exceeds a specified threshold, i.e., if

$$||\hat{\beta}_i(k|k)|| \geq \rho_i \tag{54}$$

Because the ranges of the bias magnitudes for which the sensors can function properly are dependent on the sensor type, different thresholds ρ_i may be chosen for the individual sensors. The thresholds $\{\rho_i\}$ are again determined by monitoring past sensor performance.

4.2. Failure Classification and Compensation. Because all online sensors are assumed to exhibit time-varying biases, even under nominal conditions, the classification task is reduced to discriminating between faulty and normal operation. The failure classification problem is thus completely addressed by the previously discussed failure detection scheme.

Whenever one of tests 1, 2, or 4 detects a sensor failure, the faulty sensor is excluded from contributing to the estimation process by setting the appropriate elements of the output matrix $C(k)$ to zero. Thus, the measurements of faulty sensors do not enter the estimation process.

5. Application Study: Sensor Fusion for an Ethylene Copolymerization Process

The production of ethylene copolymers of consistent composition requires tight control of the reactor feed composition. In this particular example, an ethylene copolymer grade was produced with vinyl acetate as the co-monomer. The process variable to be monitored is the feed composition, as characterized by the mole percentage of vinyl acetate. Several indicators of this variable are available:

Table 1. Summary of Sensor Characteristics

sensor number	sensor type	sampling rate	time delay	σ_i^2	biased
1	GC 1	60 s	210 s	4×10^{-3}	yes
2	GC 2	60 s	270 s	2×10^{-3}	yes
3	FTNIR	30 s	0 s	1×10^{-2}	yes
4	model	60 s	0 s	5×10^{-3}	yes
5	lab measurement	~60 min	30−90 min	2×10^{-2}	no

(1) Two gas chromatographs with sampling intervals of 60 s and measurement time delays of 210 and 270 s, respectively;

(2) A Fourier Transform Near InfraRed (FTNIR) measurement, which is available every 30 s and has no measurement time delay associated with it;

(3) A process model that provides composition estimates every 60 s; this estimate is also undelayed;

(4) In addition to these frequent indicators, samples of the polymer product are taken approximately every hour and lab analyses are performed, the results of which become available ~30−90 min after taking the samples. This laboratory measurement is considered unbiased. Therefore, it can be used as a reference measurement for sensor fusion. A summary of the sensor characteristics is given in Table 1. The measurement error variances σ_i^2 in this table were estimated from the given data record.

The first step in applying the proposed sensor fusion technique is to obtain an approximate model description of the process and the multisensor system.

Note that one of the available indicators represents a process model that could be used as a deterministic process model and augmented with a stochastic uncertainty description. Here, however, we choose to treat the model-based indicator (or "soft sensor") as a measurement device, and we use the random walk model (eq 6) to describe the process dynamics. In the current application, it is known a priori that all sensors, except for the laboratory measurement, exhibit time-varying biases. Therefore, the following sensor model is suggested, in which the sensor outputs (y_i, $i = 1, ..., 4$) are described by

$$y_i(k) = \eta(k - \theta_i) + \beta_i(k) + w_i(k) \tag{55}$$

where θ_i is a measurement time delay, w_i is random measurement noise, and β_i is a time-varying sensor bias which is again assumed to be generated by a random walk process (eq 6). The unbiased reference sensor is described by eq 7. The system, which consists of the true process variable and the measurements obtained from sensors 1−5 can be described by a single dynamic model by introducing the following state vector $x(k)$:

$$x(k) = [\eta(k) \;\; \eta(k-1) \;\; \cdots \;\; \eta(k-\theta_{\max}) \;\; \beta_1(k) \;\; \cdots \;\; \beta_4(k)]^{\mathrm{T}} \tag{56}$$

where $\theta_{\max}$ is the largest of the individual sensor time delays (except for the reference sensor). Using this state vector, the combined process and sensor model can be written in standard state-space form:

$$x(k+1) = Ax(k) + v(k) \tag{57}$$

$$y(k) = C(k)x(k) + w(k) \tag{58}$$

where $v(k) = [v_p(k) \; 0 \; \cdots \; 0 \; v_1(k) \; \cdots \; v_4(k)]^{\mathrm{T}}$ and $w(k) = [w_1(k) \; \cdots \; w_4(k)]^{\mathrm{T}}$. The covariance matrixes associated with $v(k)$ and $w(k)$ are assumed to be diagonal, i.e., $\mathrm{cov}[v(k), v(k)] = \mathrm{diag}(P_p \; 0 \; \cdots \; 0 \; P_1 \; \cdots \; P_4)$ and $\mathrm{cov}[w(k), w(k)] = \mathrm{diag}(\sigma_1^2 \; \cdots \; \sigma_5^2)$. The output matrix $C(k)$ is time-varying, because the sampling rates are not identical for all sensors. Therefore, $C(k)$ is dependent on which measurements are available at a particular sampling instant. Using the system description given as eqs 57 and 58, the Kalman filter (see eqs 21−23) provides estimates of the current state of the system, based on measurements up to time k. Because the time delay associated with the reference sensor is extremely large (>100 sampling instants), it would be impractical to incorporate the reference sensor by adding delayed versions of $\eta(k)$ as state variables. Here, we use the approach of rewinding the Kalman filter when a reference measurement becomes available, as discussed previously.

5.1. Model Parameters. Although the system matrices A and C directly follow from the dynamic characteristics of the multisensor system, there are several additional "tuning" parameters that must be specified. Although the approaches suggested here are by no means intended to be generally valid, they seemed appropriate for the case study considered, and they resulted in satisfactory performance of the sensor fusion technique.

One important guideline in determining "tuning" parameters for robust sensor fusion is to avoid relying on the results obtained from the filter in estimating these parameters. The tuning parameters considered are given as follows:

(a) The process model parameter P_p. Before implementing the Kalman filter, the only estimates of the process variable available are the measurements. Therefore, eq 51 is used to estimate P_p by replacing $\hat{\eta}(k)$ with $y_i(k)$. This is done for all sensors i, and the results are averaged.

(b) The variances P_i associated with the random sequences generating the biases. To estimate P_i, reference measurements are needed. At sampling instants at which reference measurements were taken, the sensor biases are approximated by

$$\beta_i(k) \approx y_{\mathrm{ref}}(k) - y_i(k) \tag{59}$$

and, for consecutive reference measurements obtained at samples k_1 and k_2, respectively, estimates of P_i are obtained from

$$\hat{P}_i \approx \frac{[\beta_i(k_2) - \beta_i(k_1)]^2}{k_2 - k_1} \tag{60}$$

which are, again, averaged over several reference measurements.

(c) The measurement error variances Q_i. These parameters might be given in the sensors' specifications that are provided by the sensor manufacturer, or, alternatively, the sensor variances can be estimated from measurements during time periods in which there were hardly any changes in the process variable and the bias.

This procedure is not applicable to the reference sensor. There, it is necessary to process several samples taken at the same time to estimate the sensor variance.

5.2. Results. We now discuss some results that are representative of the sensor fusion technique's performance when it was applied to actual process sensor data

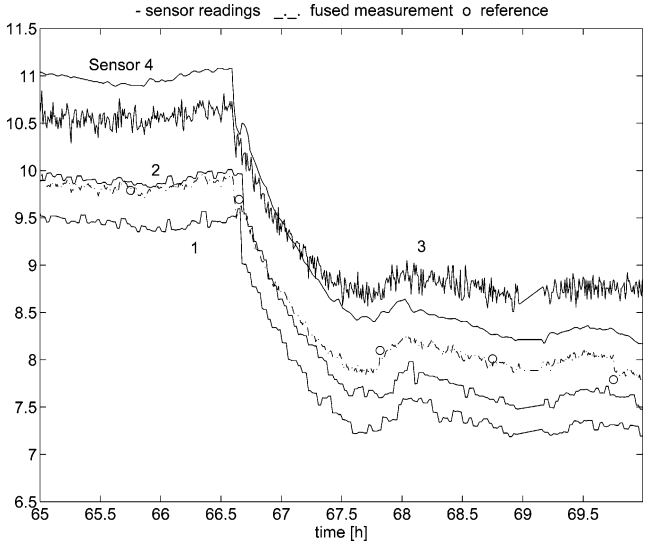

Figure 5. Sensor fusion during a process transition (data set 1).

covering a time period of eight consecutive days. The following parameters were chosen:

(a) The sampling period for the algorithm is chosen as $\Delta t = 30$ s. Hence, the time delays are $\theta_1 = 7$, $\theta_2 = 9$, $\theta_3 = \theta_4 = 0$.

(b) Additional model parameters (besides matrices A and C) required for the sensor fusion approach were determined by following the guidelines given in the previous section: The variance of the random sequence, which is assumed to generate the true process variable (see eq 6) was determined as $P_p = 0.003$. The variances of the random sequences that are assumed to generate the sensor biases were determined as $P_1 = P_2 = P_3 = 5 \times 10^{-4}$ and $P_4 = 1 \times 10^{-3}$.

(c) The window size for performing the bias and variance tests on the innovations is $N_W = 25$. The same window size is used to estimate P_p online, according to eq 51.

(d) The thresholds $\{\alpha_i\}$ for test 1 on the magnitude of the innovations are chosen as $\alpha_i = 5$, $i = 1, ..., 4$ (i.e., 5 times the theoretical standard deviation).

(e) The assumed variances of the standardized innovations required by the bias and variance tests are chosen as $\gamma_i = 1.25$ (i.e., 25% above their theoretical values of 1).

(f) The thresholds $\{\rho_i\}$ on the magnitude of bias estimates (test 4) are chosen as $\rho_i = 3$.

In the following results, we assume that the laboratory measurements are undelayed, because the actual delays of these measurements were unknown. Although the sensor fusion was performed for the entire set of sensor data, only a few representative data sequences are chosen here, to demonstrate the following: (i) the performance of the sensor fusion approach under "normal" condition; (ii) the detection and removal of measurement outliers; (iii) the detection of gradual sensor failures; and (iv) the detection of "abnormal" process behavior.

5.2.1. Data Set 1. Here, we examine the performance of the sensor fusion scheme during transitions between two different polymer products. Figure 5 shows that the process transition seems to be followed quite accurately by the fused measurement. Again, the model parameter P_p is estimated online (see Figure 6). During the initial phase of the transition, $\hat{P}_p(k)$ exceeds its nominal value, and the estimates are used to update the process

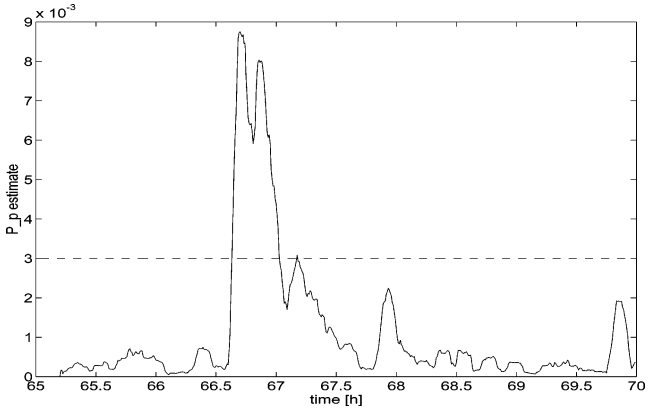

Figure 6. On-line estimation of the model parameter P_p (data set 1).

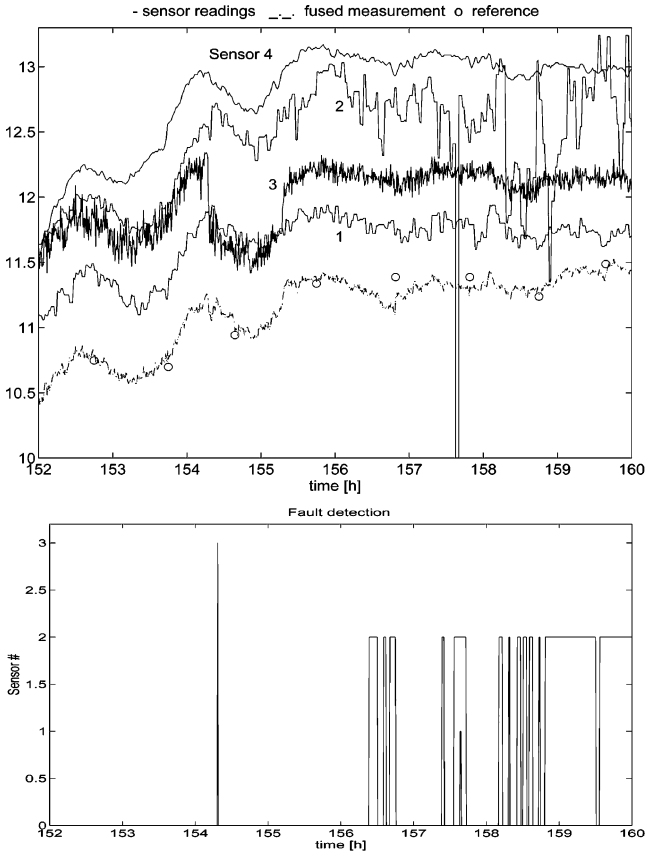

Figure 7. Sensor fusion for data set 2.

description. The online estimation of P_p seems to be capable of tracking the changes in the process behavior, as indicated by the fact that the failure detection scheme did not trigger a sensor failure in any of the sensors.

5.2.2. Data Set 2. Figure 7 illustrates the detection of a sudden shift in one sensor and a gradual failure of another sensor. The combined measurement does not seem to be affected by the severe degradation of sensor 2, nor is it affected by the sudden shift in sensor 3 at $t \approx 154.5$, which was detected by the outlier test (test 1). Test 2 was capable of detecting an imminent failure of sensor 2 very promptly, because the measurement error variance of sensor 2 exhibited a significant increase starting at a value of $t \approx 156$.

5.2.3. Data Set 3. This data set contains multiple sensor failures. The performance of the sensor fusion and fault detection schemes is shown in Figure 8. The lower trace of the figure shows the decisions of the fault

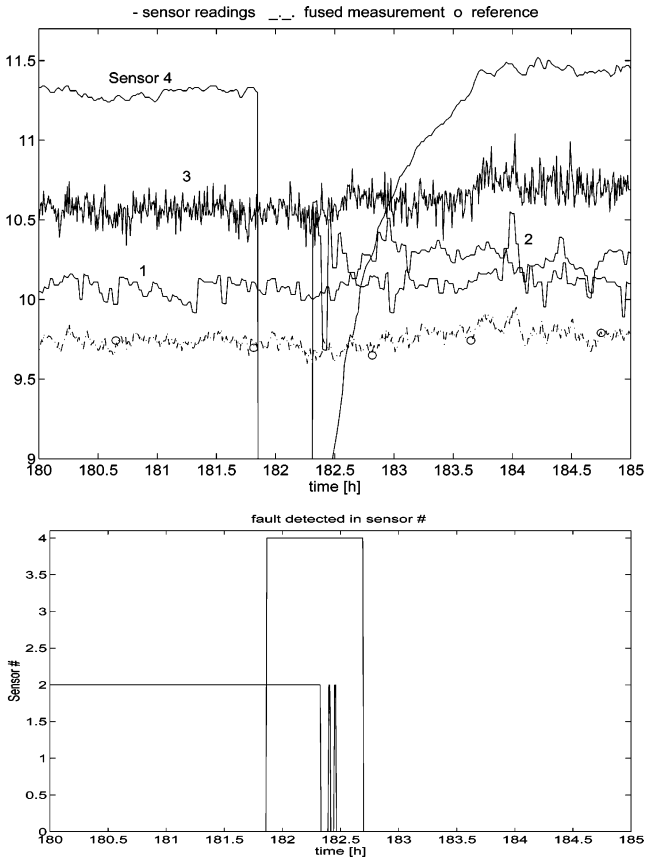

Figure 8. Sensor fusion for data set 3.

detection scheme. Note that the plot only shows the overall decisions. However, it is interesting to indicate which tests triggered at a particular time. At the beginning of this data set, sensor 2 is detected as being faulty, as indicated by test 4 (test on the bias magnitude). Shortly before $t = 182$, sensor 4 fails, which is detected initially by test 1, and then by test 4. Because the magnitude of the bias estimate remains above the threshold of test 4 (ρ_4) for some time, test 4 continues to be triggered until $t \approx 182.2$. After that, a sensor failure of sensor 4 continues to be indicated by test 2, because the bias is changing rapidly, resulting in an inflation of the variance of the corresponding innovation sequence.

6. Summary and Conclusions

We have presented a sensor fusion approach that has been developed using stochastic system theory but inspired by experimental evidence of integrated fusion of all available disparate sensory inputs in the control function performed by baroreflexes, which are the biological systems responsible for short-term blood-pressure control in mammals. More specifically, we addressed the problem of how to combine measurements of a process variable that are available from different sources into one single, robust, and more-precise estimate; moreover, we have done so within the industrial context where the unavoidable imperfections and stringent requirements on robustness to sensor bias and other partial sensor failures must be addressed explicitly.

We derived algorithms for obtaining optimum estimates of the true process variable that are more precise, more accurate, and more failure tolerant than that obtainable from each individual sensor. The advantages of utilizing multiple redundant sensors, as well as delayed sensors and multirate sensors, have been quantified. Robustness was achieved by augmenting the nominal technique with failure detection, classification, and compensation techniques, based on statistical hypothesis testing. The performance of the proposed techniques was demonstrated via an actual application on an industrial ethylene copolymerization process.

The primary limitation of the technique is related to its direct connection with the Kalman filter. Even though we explicitly discussed how to make the algorithm robust to the key idealized assumptions underlying the Kalman filter, it is not clear how the technique will perform on a severely nonlinear process. However, the application on the industrial ethylene copolymer process provides confirmation that the technique will, in fact, function well on actual industrial processes, and it has been tested during grade transitions when process nonlinearities can be expected to be present.

Although the entire development was inspired by the baroreceptor reflex, not much biological detail of the blood-pressure control system is contained in the algorithm. It would be interesting, as more-detailed knowledge of the biological system's sensor fusion mechanism becomes available, to revisit the problem and develop a truly biomimetic algorithm that faithfully mimics "in silico" the biological sensor fusion mechanism with significant biological fidelity.

Nevertheless, even in its current stochastic systems theory manifestation, the technique explicitly addresses only the analysis problem: i.e., given the multisensor configuration, it produces the optimal fused estimate. A possible next step is to investigate how this framework can be used for the design problem: i.e., given a specified objective of system precision and robustness, determine the number, characteristics, and configuration of the sensor system required to meet the objective.

Appendix A. Development of the Statistical Tests

A.1. Bias Test. Several failure detection mechanisms described in this article are used to perform statistical tests on the recent history of random variables. First, we develop a statistical test that determines whether a sample (which, in this case, corresponds to a sequence of consecutive observations) has a zero or nonzero mean. This test will be performed for data sequences of increasing length, starting with data corresponding to the current sampling instant and successively including more past data until a decision can be made with a certain confidence or until the entire data set (of length N_W) has been analyzed. Such a sequential procedure ensures that biases are detected as soon as possible.

Before describing the sequential test implementation, we discuss the basic tests for data sequences of constant size N. Let us assume for the moment that the elements of the set X (for example, x_i) are normally distributed random variables with an unknown mean μ and unity variance $\mathcal{N}(\mu, 1)$. As will be seen later, this assumption of normality is not required to apply the bias test. However, we will derive this test using the normality assumption. The following two tests are performed in parallel.

A.1.1. Test 1. This test is designed to detect a nonzero mean of the data sequence X. The hypothesis that the mean of the sequence is identical to zero (H_0: $\mu = 0$) is tested versus the alternative hypothesis that the mean

is unequal zero (H_1: $\mu \neq 0$). The likelihood ratio for a data sequence of length N is[33]

$$\frac{L(0; x_1, x_2, ..., x_N)}{L(\hat{\mu}; x_1, x_2, ..., x_N)} = \prod_{i=1}^{N} \frac{p_1(x_i)}{p_0(x_i)} = \exp\left[-\frac{\sum_{i=1}^{N} x_i^2}{2} + \frac{\sum_{i=1}^{N} (x_i - \hat{\mu})^2}{2}\right] \quad (61)$$

where p_0 and p_1 are Gaussian probability densities that correspond to hypotheses H_0 and H_1, and $\hat{\mu}$ is the maximum likelihood estimate of the new mean, $\hat{\mu} = \sum_{i=1}^{N} x_i/N$. After simplifying eq 61 and noting that, under hypothesis H_0, the random variable

$$\frac{\sum_{i=1}^{N} x_i}{\sqrt{N}} \sim \mathcal{N}(0, 1) \quad (62)$$

the likelihood ratio test can be stated as follows:

$$\text{accept } H_1 \text{ if } \frac{|\sum_{i=1}^{N} x_i|}{\sqrt{N}} > c_1; \text{ do not accept } H_1 \text{ otherwise} \quad (63)$$

Here, c_1 is chosen to result in a specified probability of a Type I error (accepting H_1 when H_0 is true). Note that this test is capable of detecting changes in the mean of the sequence X with a specified confidence. However, the test does not provide a measure of confidence in hypothesis H_0 if H_1 has not been accepted. In this case, the test simply concludes that there is not sufficient evidence to accept H_1; however, H_0 may be true or not. Therefore, we apply a second test as follows.

A.1.2. Test 2. This test is designed to detect zero-mean sequences. The hypothesis that the absolute value of the mean is larger than a specified minimum value (H_2: $|\mu| > \mu_{min}$) is tested versus the hypothesis that the sequence has a zero mean (H_0: $\mu = 0$). A minimum nonzero bias (μ_{min}) must be specified in this case, to enable the rejection of H_0, given a finite data set. Noting that, under hypothesis H_2, the random variable

$$\frac{\sum_{i=1}^{N} x_i}{\sqrt{N}} \sim \mathcal{N}(\mu, 1) \quad (64)$$

the following test can be stated:

$$\text{accept } H_0 \text{ if } \frac{|\sum_{i=1}^{N} x_i|}{\sqrt{N}} < \mu_{min}\sqrt{N} - c_2; \text{ do not accept } H_0 \text{ otherwise} \quad (65)$$

Again, c_2 is chosen to yield a desired significance level of the test. The parameters of tests 1 and 2 must be selected appropriately, such that contradictory decisions (hypotheses H_0 and H_1 are accepted simultaneously) are avoided. Hence, the parameters c_1, c_2, μ_{min}, and N must satisfy the relation

$$c_1 \geq \mu_{min}\sqrt{N} - c_2 \qquad (\forall N) \quad (66)$$

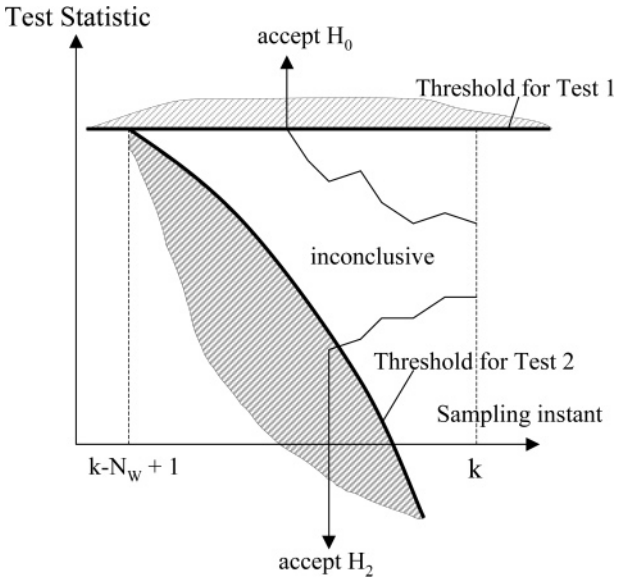

Figure 9. Illustration of sequential hypothesis testing (bias test).

Because the threshold for test 1 is constant and the threshold for test 2 is monotonically increasing with N, it is sufficient that eq 66 holds for the maximum window size N_W. If eq 66 holds as an equality at the maximum window size N_W, a conclusive decision (i.e., either H_0 or H_1 is accepted) is guaranteed.

By inspecting eqs 62 and 64, we notice that the statistical tests described here should be applicable to any type of random sequences, as long as $\sum_{i=1}^{N} x_i/\sqrt{N}$ is *close* to the normal distribution indicated in eq 62 (or eq 64). Now, it follows from the central limit theorem (see, e.g., ref 25, p 175) that, for any sequence of independent, identically distributed random variables with finite means and unity variance, but irrespective of the actual distributions, the statistics given in eqs 62 and 64 approach the corresponding normal distributions for $N \rightarrow \infty$. Therefore, from a practical point of view, the bias test can be applied, even if the random variables involved follow distributions that are quite different from the normal distribution, provided that large window sizes N are used.

The tests are applied in a sequential fashion to the sequence X and the parameters of the statistical tests are chosen such that eq 66 is satisfied as an equality for $N = N_W$. The sequential test starts with a window size of $N = 1$ and performs tests 1 and 2. If neither test accepts its alternative hypothesis (H_1 and H_0, respectively), the number of values used in the data sequence is increased by one (which corresponds to including more past data) and the tests are performed again. The window size N is increased until one of the tests accepts its alternative hypothesis. By design, this must occur at the maximum window size N_W, at the latest. The implementation of the statistical tests is graphically illustrated in Figure 9.

Note that, in the sequential implementation of tests 1 and 2, several hypothesis tests are typically performed before a decision is made. Therefore, the probability of a Type I error in the sequential implementation will be larger than the value specified for a single hypothesis test. Using $c_1 = c_2$, these limits have been adjusted for

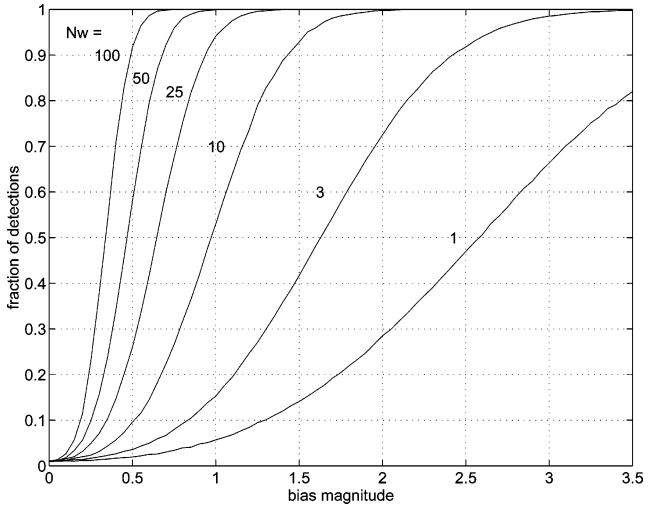

Figure 10. Performance of bias test (Type I error probability = 1%).

Table 2. Thresholds for Sequential Bias Test with Type I Error Probabilities of 5%, 2.5%, and 1%

	threshold $c_1 = c_2$		
window size, N_W	5%	2.5%	1%
1	1.96	2.24	2.58
3	2.25	2.54	2.82
10	2.51	2.78	3.09
25	2.68	2.93	3.24
50	2.76	3.02	3.32
100	2.84	3.10	3.40

window sizes of $N > 1$, such that the probability of a Type I error (rejecting H_0, although it is true) in the sequential test implementation remains constant at a desired level. These thresholds for various Type I error probabilities (5%, 2.5%, 1%) are summarized in Table 2.

Monte Carlo simulations are performed to evaluate the performance of the proposed sequential test. Figure 10 shows the results of such simulations for a Type I error probability of 1%. The data were drawn randomly from $\mathcal{N}(1, 0)$, and biases of increasing magnitudes have been added to the data sequences. The percentage of decisions where the tests detected bias in the given data sequence is plotted versus the actual bias in the data. As can be seen, the performance of the tests is dependent on the maximum window size N_W.

A.2. Variance Test. Here, we intend to derive a test that will detect an increase in the variance of a random variable. The basic test used is the following.

A.2.1. Test 3. Under the hypothesis H_0, that the sequence of observations of a random variable X follows a normal distribution with known variance σ^2 but unknown mean, the statistic

$$\Gamma(N) = \sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{\sigma^2} \tag{67}$$

where

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

follows a central χ^2 distribution, with $N - 1$ degrees of freedom ($\chi^2(N - 1, 0)$).[33] Under the alternative hypothesis H_1—that the variance is, in fact, larger than the assumed value—the statistic $\Gamma(N)$ would be inflated. This suggests the following test:

$$\text{reject } H_0 \text{ if } \Gamma(N) > c(N) \tag{68}$$

where the threshold $c(N)$ is typically chosen by specifying a desired Type I error probability. This test is strongly dependent on the normality assumption and, therefore, is somewhat "risky" to apply in practice, where deviations from normality are very common. In particular, both strong asymmetries and outliers are likely to cause difficulties for such a test. Although we do not provide comprehensive measures that would avoid problems with this test that result from any deviations from normality, it needs to be noted that, within the sensor fusion techniques described in this article, this variance test is not applied to "raw process data", but only to data that have been cleaned of outliers. Therefore, the test can be expected to perform well, even in the presence of certain types of non-normality that are observed in practice.

Table 3. Correction Factors f_2 for the Thresholds of the Sequential Variance Test for Type I Error Probabilities of 5%, 2.5%, and 1%

	f_2		
window size, N_W	5%	2.5%	1%
2	1.0	1.0	1.0
5	1.26	1.23	1.17
10	1.33	1.29	1.24
25	1.37	1.32	1.27
50	1.38	1.325	1.275
100	1.39	1.33	1.28

To detect changes in the variance quickly, the test is performed in a sequential fashion for data sequences of increasing length (i.e., data sets that include more and more past information). Similar to the sequential test for bias detection, the test is implemented in the following fashion. Starting with the statistic at the two most recent sampling instants (i.e., $N = 2$), the window size N is increased until hypothesis H_0 is rejected. If H_0 is not rejected at the maximum window size N_W, then H_0 is accepted.

The thresholds $c(N)$ are determined as follows: First, a sequence $c_0(N)$, $2 \le N \le N_W$ is determined from the percentage points of the appropriate χ^2 distribution, by specifying a desired Type I error probability for test 3 (rejecting H_0 when H_0 is true). However, as has been noted previously, the Type I error probabilities in the sequential test implementation do not coincide with those for a data set of fixed size. Hence, the thresholds $c_0(N)$ must be adjusted, such that the Type I error probability in the sequential test implementation equals a certain desired value. This adjustment is performed by increasing the entire sequence of thresholds $c(N)$ by a constant factor, such that the overall test achieves the desired Type I error probability:

$$c(N) = f_2 c_0(N) \qquad (\text{for } N = 1, \ldots, N_W) \tag{69}$$

The constant f_2 values have been determined numerically via Monte Carlo simulations; they are summarized in Table 3.

Nomenclature

A, B, C, D, F = system matrices
f_1, f_2 = multipication factors
H = hypothesis
k = sampling instant

K = Kalman filter gain
M = period of multirate sensor
n = number of sensors
N = length of a data sequence
N_W = window size (for statistical tests)
P = covariance matrix associated with v
Q = covariance matrix associated with w
T = a posteriori estimation error covariance
v = random disturbance
w = random measurement noise
x = state vector
x_i = state vector associated with sensor i
$\hat{x}$ = state estimate
ψ = measurement vector
α_i = threshold for outlier test
β_i = sensor bias
γ_i = assumed variance of innovation sequence
Δt = sampling period
ϵ = estimation error
η = true process variable
$\hat{\eta}$ = estimated process variable (using all available sensors)
$\hat{\eta}_i$ = estimated process variable (using sensor i only)
θ = time delay (in multiples of Δt)
μ = mean
ν = threshold for sensor consistency
ρ_i = threshold for bias magnitude
σ = standard deviation
Σ = a priori estimation error covariance
τ = time constant
ϕ = efficiency factor
ω = innovation
$\tilde{\omega}$ = standardized innovation

Subscripts

c = combined
i = associated with sensor i
p = process
ref = reference sensor
s = sensor system

Acknowledgment

We thank Cliff Phares (DuPont Packaging & Industrial Polymers) for providing the process data for section 5, for many discussions on how to make the approach work in practice, and for his assistance in making possible the first online implementation of the technique.

Literature Cited

(1) Luo, R. C.; Kay, M. G. Multisensor Integration and Fusion in Intelligent Systems. *IEEE Trans. Syst., Man Cybernet.* **1989**, *19*, 901−931.

(2) Crowley, J. L.; Demazeau, Y. Principles and Techniques of Sensor Data Fusion. *Signal Process.* **1993**, *32*, 5−27.

(3) Kerr, T. Decentralized Filtering and Redundancy Management for Multisensor Navigation. *IEEE Trans. Aerosp. Electron. Syst.* **1987**, *23*, 83−119.

(4) Dasarathy, B. V. Decision Fusion Strategies in Multisensor Environments. *IEEE Trans. Syst., Man Cybernet.* **1991**, *21*, 1140−1154.

(5) Mitchell, A.; Willis, M. J.; Tham, M. T.; Johnson, N.; Skorge, L. M. Inferential Estimation of Viscosity Index on a Lubricant Production Plant. In *Proceedings of the American Control Conference*, Seattle, WA, 1995; IEEE: Piscataway, NJ, 1995; pp 24−28.

(6) Cimander, C.; Carlsson, M.; Mandenius, C.-F. Sensor Fusion for On-Line Monitoring of Yoghurt Fermentation. *J. Biotechnol.* **2002**, *99*, 237−248.

(7) Ramamurthy, Y.; Sistu, P. B.; Bequette, B. W. Control-Relevant Dynamic Data Reconciliation and Parameter Estimation. *Comput. Chem. Eng.* **1993**, *17*, 41−59.

(8) Tjoa, I. B.; Biegler, L. T. Simultaneous Strategies for Data Reconciliation and Gross Error Detection in Nonlinear Systems. *Comput. Chem. Eng.* **1991**, *15*, 679−690.

(9) Negiz, A.; Cinar, A. Monitoring of Multivariable Dynamic Processes and Sensor Auditing. *J. Process Control* **1998**, *8*, 375−380.

(10) Sasiadek, J. Z. Sensor Fusion. *Ann. Rev. Control* **2002**, *26*, 203−228.

(11) Willsky, A. S. A Survey of Design Methods for Failure Detection in Dynamic Systems. *Automatica* **1976**, *12*, 601−611.

(12) Takahashi, K.; Nozaki, S. From Intelligent Sensors to Fuzzy Sensors. *Sens. Actuators A* **1994**, *40*, 89−91.

(13) Hong, L.; Wang, G.-J. Centralised Integration of Multisensor Noisy and Fuzzy Data. *IEE Proc.—D: Control Theory Appl.* **1995**, *142*, 459−465.

(14) Lee, S. C. Sensor Value Validation Based on Systematic Exploration of the Sensor Redundancy for Fault Diagnosis KBS. *IEEE Trans. Syst., Man Cybernet.* **1994**, *24*, 594−605.

(15) Naidu, S. R.; Zafiriou, E.; McAvoy, T. J. Use of Neural Networks for Sensor Failure Detection in a Control System. *IEEE Control Syst. Mag.* **1990**, *10* (4), 49−55.

(16) Da, R.; Lin, C.-F. Sensor Failure Detection with a Bank of Kalman Filters. In *Proceedings of the American Control Conference*, Seattle, WA, 1995; IEEE: Piscataway, NJ, 1995; pp 1122−1126.

(17) Sun S.-L.; Deng, Z.-L. Multi-Sensor Optimal Information Fusion Kalman Filter. *Automatica* **2004**, *40*, 1017−1023.

(18) Faggin, F.; Mead, C. VLSI Implementation of Neural Networks. In *An Introduction to Neural and Electronic Networks*; Academic Press: San Diego, CA, 1990; pp 275−292.

(19) Henson, M. A.; Ogunnaike, B. A.; Schwaber, J. S.; Doyle, F. J. The Baroreceptor Reflex: A Biological Control System with Applications in Chemical Process Control. *Ind. Eng. Chem. Res.* **1994**, *33*, 2453−2466.

(20) Pottmann, M.; Henson, M. A.; Ogunnaike, B. A.; Schwaber, J. S. A Parallel Control Strategy Abstracted from the Baroreceptor Reflex. *Chem. Eng. Sci.* **1996**, *51*, 931−945.

(21) Seagard, J. L.; van Brederode, J. F. M.; Hopp, F. A.; Dean, C.; Gallenburg, L. A.; Kampine, J. P. Firing Characteristics of Single-Fiber Carotid Sinus Baroreceptors. *Circulat. Res.* **1992**, *66*, 1499−1509.

(22) Karemaker, J. M. Neurophysiology of the Baroreceptor Reflex. In *The Beat-by-Beat Investigation of Cardiovascular Function*; Clarendon Press: Oxford, U.K., 1987; pp 27−49.

(23) Rogers, R. F. Arterial Baroreceptor Signal Processing in the Nucleus Tractus Solitarius, Ph.D. Dissertation, University of Pennsylvania, Philadelphia, PA, 1995.

(24) Rogers, R. F.; Rose, W. C.; Schwaber, J. S. Encoding of Carotid Sinus Pressure and dP/dt by Myelinated Baroreceptor Target Neurons in the Nucleus Tractus Solitarius. *J. Neurophysiol.* **1996**, *76*, 2644−2660.

(25) Grimmett, G. R.; Stirzaker, D. R. *Probability and Random Processes*, 2nd ed.; Clarendon Press: Oxford, U.K., 1992.

(26) Gelb, A. *Applied Optimal Estimation*, MIT Press: Cambridge, MA, 1980.

(27) Tyler, M. L.; Morari, M. Estimation of Cross-Directional Properties: Scanning vs Stationary Sensors. *AIChE J.* **1995**, *41*, 846−854.

(28) Goodwin, G. C.; Sin, K. S. *Adaptive Filtering, Prediction, and Control*; Prentice Hall: Englewood Cliffs, NJ, 1984.

(29) Zambare, N.; Soroush, M.; Ogunnaike, B. A. A Method of Robust Multi-rate State Estimation. *J. Process Control* **2003**, *13*, 337−355.

(30) Rawlings, J. B.; Chien, I.-L. Gage Control of Film and Sheet Forming Processes. *AIChE J.* **1996**, *42*, 753−766.

(31) Campbell, J. C.; Rawlings, J. B. Predictive Control of Sheet- and Film-Forming Processes. *AIChE J.* **1998**, *44*, 1713−1723.

(32) Mehra, R. K.; Peschon, J. An Innovations Approach to Fault Detection and Diagnosis in Dynamic Systems. *Automatica* **1971**, *7*, 637−640.

(33) Hogg, R. V.; Craig, A. T. *Introduction to Mathematical Statistics*, 4th ed.; Macmillan: New York, 1978.

Received for review May 10, 2004
Revised manuscript received August 23, 2004
Accepted August 30, 2004

IE049614T