Adaptive Nonlinear Cell Mass State Estimator for a Continuous Yeast Fermentation

Michael E. Ramseier, Pramod Agrawal, and Duncan A. Mellichamp
Dept. of Chemical and Nuclear Engineering, University of California, Santa Barbara, CA 93106

On-line measurement of the important state variables in fermentations, particularly cell mass concentration, remains a difficult problem. However, a number of secondary or environmental variables can be measured conventionally and on-line, such as pH, and CO_2 and O_2 in the exhaust gas. Stephanopoulos and San (1984) have developed a modeling approach, based on species balances, that provides relations between the environmental and important state variables. Using such a model, the important state variables can be estimated in principle from more easily accessible on-line measurements.

In this article, a new adaptive estimator is developed, incorporating as its basis an underlying nonlinear model so as to utilize the best possible a priori process knowledge. Base addition rate and CO_2 offgas concentration are measured on-line and periodically. Cell mass measurements are incorporated infrequently and even at irregular sampling periods, thus providing a very flexible scheme. Only a single adapted parameter is required to match the model to the plant operating characteristics. This simple but rigorous model form results in an estimator that is easy to implement and to tune and which exhibits long-term robustness due to its multirate feedback structure. Experimental results from a laboratory-scale continuous fermentor show that such a cell mass estimation scheme yields excellent performance both open-loop (without control) and as a part of conventional and nonlinear adaptive control approaches.

Introduction

Knowledge of fermentation process state variables is crucial in any type of control scheme. Important process state variables must be available to obtain tight control of a system at its optimal operating conditions. Unfortunately, with fermentations the important process state variables—cell mass, substrate, and product concentrations—are not easily measured. On-line measurement of these variables particularly remains a difficult problem. On-line gas chromatographs and high performance liquid chromatographs can be used; these, however, require expensive sampling devices and create long measurement time delays which ultimately deteriorate the overall control performance. More promising is the use of so-called biosensors (involving enzymes). Halme (1987) reviews recent developments in the field of biosensors, but also discusses their potential shortcomings.

Correspondence concerning this article should be addressed to D. A. Mellichamp.

Alternative approaches to direct measurement are available. For example, a number of variables can be measured conventionally on-line, such as pH, temperature, CO_2 and O_2 in the exhaust gas, and dissolved oxygen. It has been suggested that only these environmental variables be controlled directly in order to maintain the primary state variables indirectly. Agrawal (1989) studied an acid production rate controlled fermentation, where the acid production rate is obtained by controlling the pH at a constant value and measuring the amount of base added. Chang and Lim (1990) suggested a similar approach, where the CO_2 evolution rate is used to optimize the cellular productivity. Considering the complexity of cell growth mechanisms, it is questionable if such methods can actually keep the unmeasured and important liquid concentrations at their desired levels. Another approach is to model the relationships between environmental and important state variables, and then to infer the states from easily accessible measurements. Ste-

phanopoulos and San (1984) discuss in detail how such models based on species balances can be developed. In recent years several authors (for example, Bastin and Dochain, 1990) have concluded that state estimation is the most promising technique, and much effort is underway to develop and to improve such estimation schemes, including the work described below.

Successful application of Kalman filters for state estimation in fermentation processes has been reported in the literature. Stephanopoulos and San (1984) and Bellgardt et al. (1986) discuss experimental results of the application of an extended Kalman filter (EKF) to fed-batch and batch fermentations, respectively. The EKF structure is well suited for fermentation systems because of its capability to deal with nonlinear process relations and to provide combined state and parameter estimation, which can be very attractive. For given model uncertainty and measurement noise characteristics, the EKF gives the optimal state estimate of a stochastic system. A general treatment of the Kalman filter theory can be found in Ljung (1979) or Goodwin and Sin (1984).

However, several general comments concerning the application of the EKF theory to fermentation systems should be made. Model linearization around the most recent estimate of the state variables can lead to numerical problems and convergence difficulties, as reported by Ljung (1979). The EKF gives optimal estimates only if model uncertainty and measurement noise characteristics are known exactly. This requirement is not easily satisfied with fermentation systems where, for example, it is hard to quantify the modeling error. Thus, Stephanopoulos and San (1984) assumed no modeling error, and Ramirez (1987) set the model uncertainty equal to the measurement noise. Agarwal and Bonvin (1989) recently have elaborated on the importance of the proper choice of modeling error variances.

To estimate model parameters along with the state variables (as in the EKF), one has to assume certain dynamics for the parameters because they are treated in the same way as the state variables. The selection of such parameter dynamics typically is done in a rather *ad hoc* fashion, and this approach can lead to better or worse parameter estimates as has been discussed in De Vallière and Bonvin (1990). Furthermore, the coupling of parameter and state estimation can cause an overall performance deterioration. In summary, the EKF algorithm requires a large design effort and *a priori* estimates of important noise and uncertainty characteristics.

The state variable estimation methods applied to fermentation processes that have been discussed to this point use only environmental variables as measurements. Such an approach deviates from the conventional EKF algorithm, where the process output variables are measured directly, and potentially can lead to drifting state estimates (offsets) since the model that relates state and environmental variables can be inaccurate or slowly time-varying. This issue is not so important for batch or fed-batch operations which usually exhibit fermentation times of less than one or two days. Continuous fermentations or extended fed-batch fermentations lasting several days, however, require long-term robust state estimators. The only way to check if a state estimator exhibits offset or not, is by measuring the state variables off-line. This check will usually be made only infrequently, while on-line measurements are taken more often. It is possible to combine the two forms of measurements within the same estimator, using a multirate sampling scheme. Lewis (1986) describes such a scheme applied to a Kalman filter. The idea behind this approach is to utilize frequently accessible on-line measurements to obtain reasonable state estimates continuously. Since these on-line measurements are not of the actual state variables, this estimation technique has an open-loop structure. The estimation loop can be closed, however, by feeding infrequent off-line measurements of the state variables into the estimator.

Instead of applying this methodology to a Kalman filter, Guilandoust et al. (1988) and Tham et al. (1988) utilized a different approach. They formulated a linear black-box model with unknown parameters, relating the state variables to infrequent off-line and frequent on-line measurements. The parameters of this model are updated whenever a new off-line measurement is available. In between off-line samples, these parameters are used along with the on-line measurements to estimate the states. The authors referred to their scheme as an adaptive estimator.

The concept of an adaptive estimator is attractive because it makes use of easily available *on-line* process information yet ensures long-term robustness by incorporating high-level *off-line* information at a much slower sampling rate. The state estimator developed in this article incorporates a similar adaptive structure; thus it avoids both the critical model linearization step used by the EKF and the need for *a priori* knowledge of model uncertainty. The key element here is the use of a nonlinear process model which should yield better estimates than an estimator employing a linear black-box model. Furthermore, as described below, it is possible to reduce the number of adapted model parameters to two by incorporating all available *a priori* process knowledge. One of the two decoupled parameter estimators can be further reduced to a simple algebraic equation.

This work was confined to a single state variable estimation, the cell mass concentration. The substrate concentration in a substrate-limited continuous yeast culture is usually quite low (less than 0.1 g/L) and is therefore difficult to obtain reliably from off-line measurements with standard laboratory equipment. The product concentration could be measured off-line, but is only of limited use if the substrate concentration is not available. A MIMO (multiple-input/multiple-output) control system requires the availability of all three state variables. However, the approach developed here for the cell mass concentration estimator can be easily modified and incorporated in substrate and product estimators, if both variables can be reliably measured off-line.

The following sections describe in more detail how the adaptive nonlinear estimator was designed and implemented.

Development of Cell Estimator Model

In order to use environmental variables to estimate the unmeasured state variables, one first has to model their interrelationship. Stephanopoulos and San (1984) first presented a modeling methodology which now is applied by many others. Here, we use the same approach, but modify the model to match the process characteristics of the particular experimental fermentation system.

The model assumes an overall chemical reaction for the cell growth in a synthetic medium containing ammonium ions as the only source of nitrogen. It takes into account all major participating species:

$$\nu_1 C_6H_{12}O_6 + \nu_2 O_2 + \nu_3 (NH_4)_2SO_4 \rightarrow C_{\gamma_1}H_{\gamma_2}O_{\gamma_3}N_{\gamma_4} + \nu_4 H_2O + \nu_5 CO_2 + \nu_6 C_2H_6O + \nu_3 H_2SO_4 \quad (1)$$

sugar, oxygen, ammonium sulfate, cell mass, water, carbon dioxide, ethanol, sulfuric acid

The ν_i are the stoichiometric coefficients. An elemental laboratory analysis determines the cell mass composition coefficients ($\gamma_1 - \gamma_4$). Dried cells from this work were analyzed by Galbraith Laboratories (Knoxville, TN) with the following result:

overall composition: $CH_{1.748}O_{0.510}N_{0.156}$

formula weight: $FW_x = 24.11$ g/mol

The elemental analysis also determined that the four elements C, H, O, and N only account for 90.85% of the total cell mass in our case. The remaining cell material is referred to as ash. Because we are interested in the total cell mass, subsequent model equations have to take into account this factor of 0.9085.

The six stoichiometric coefficients can be determined by writing four atomic balances for the elements C, H, O, and N. The missing two relationships can be obtained by equating CO_2 and O_2 production rates to the cell growth rate. This approach requires that CO_2 and O_2 be measured in the gas phase. Once all six stoichiometric coefficients are known, then the growth rate of cell mass, substrate consumption rate and product formation rate are available from Eq. 1 and can be used in appropriate mass conservation equations. The derivation simplifies considerably, if only an expression for the cell mass growth rate is desired. This simplified case is presented here in detail.

Yeast cells take up molecular ammonia as their nitrogen source. Since the medium contains an ammonium salt, 1 mol of H^+ is released for each mole of ammonia that is taken up ($NH_4^+ \rightarrow NH_3 + H^+$). This H^+ production rate can be directly measured on-line by adding the equivalent amount of a strong base in order to keep the pH constant. Equation 1 basically defines the ratio between cell mass growth rate and base addition rate, and with some further manipulations, we arrive at the following model-based estimate:

$$\dot{\hat{x}}_{\text{growth}} = \dot{B}_x \frac{FW_x}{0.9085\, V_L \gamma_4} \quad (2)$$

where $\dot{\hat{x}}_{\text{growth}}$ denotes the estimate of the (total) cell mass growth rate in g/(L·h), $\dot{B}_x$ the base addition rate responsible for cell mass growth in mol/h, and V_L the fermentor liquid volume in liters. The cell mass growth rate is related to the cell mass concentration through the cell mass balance around the fermentor:

$$\dot{x} = -Dx + \mu x \quad (3a)$$

$$= -Dx + \dot{x}_{\text{growth}} \quad (3b)$$

The second form of the equation follows directly from the definition of the specific growth rate μ. The only modeling assumptions made in deriving Eqs. 3 are perfect mixing, and no cell growth in the foam layer and on the fermentor wall.

Since an estimate of the cell mass growth rate is available from Eq. 2, it is now possible to use Eq. 3b to estimate the cell mass concentration

$$\dot{\hat{x}} = -D\hat{x} + \dot{\hat{x}}_{\text{growth}} \quad (4)$$

where the caret (^) denotes an estimated variable. Thus, Eqs. 2 and 4 furnish the foundations for the nonadaptive cell mass estimator; however, further modifications are required in order to arrive at a final form which can be practically implemented. To simplify notation in the sequel, we define the cell mass growth rate as Y. Its estimate, then is:

$$\hat{Y} = \dot{\hat{x}}_{\text{growth}} \quad (5a)$$

$$= \hat{\mu}\hat{x} \quad (5b)$$

Applying the Euler approximation for the time derivative in Eq. 4, yields the discrete-time form

$$\hat{x}(t+1) = [1 - \Delta t_{\text{est}} D(t)]\hat{x}(t) + \Delta t_{\text{est}} \hat{Y}(t) \quad (6)$$

If the sampling time Δt_{est} is chosen small compared to the dominant system time constant, then the above approximation is reasonable. For example, with the experimental system a sampling time of three minutes was chosen. This value is at least a hundred times smaller than a typical fermentation time constant. It should be noted that the underlying estimation model, Eq. 6, is nonlinear but does not have to be linearized as in an extended Kalman filter scheme.

The base addition rate that is directly monitored from the process, $\dot{B}_m$, unfortunately does not coincide with $\dot{B}_x$ used in Eq. 2. There are other sources of H^+ production that are compensated as part of the measured base addition rate. One of these is the medium feed which in this case has a pH of 4.4. A certain amount of base has to be added therefore to bring the feed up to the desired fermentation pH of 5.5. $\dot{B}_f$ is used to denote the portion of the base addition rate (mol/h) that is used to compensate the low feed pH. It can be expressed as

$$\dot{B}_f = 8.26 \times 10^{-4} D\, V_L \quad (7)$$

where the constant was determined experimentally.

The second source of H^+ production is associated with the CO_2 production that accompanies cell growth. Koshy (1986) already has derived an appropriate expression for the case of a fed-batch fermentation. Following Koshy's methodology, a similar expression can be derived for a continuous fermentor, assuming that at a pH of 5.5 negligible amounts of carbonate ions are formed

$$\dot{B}_c = Q_L K C_g \quad (8)$$

$\dot{B}_c$ is the portion of the base addition rate (mol/h) that is used to compensate for the CO_2 production. Q_L is the liquid flow rate through the fermentor in L/h ($D = Q_L/V_L$), C_g is the concentration of CO_2 in the exhaust gas in mol/m^3, and K is a gas/liquid equilibrium constant. It is assumed that the gas phase dynamics are very fast compared to the cell growth dynamics, and that Eq. 8 therefore always holds. The equi-

librium constant K will vary from system to system and has to be determined experimentally.

Using the relationship

$$\dot{B}_x = \dot{B}_m - \dot{B}_c - \dot{B}_f \tag{9}$$

it is now possible to estimate the cell mass growth rate as a function of directly measurable variables as follows

$$\hat{Y} = a\left(\frac{\dot{B}_m}{V_L} - bC_gD - 8.26\cdot 10^{-4}D\right) \tag{10}$$

$$= \hat{\mu}\hat{x} \tag{5b}$$

The two parameters are defined as

$$a = \frac{FW_x}{0.9085\,\gamma_4} \quad \text{and} \quad b = K \tag{11}$$

The contribution of $\dot{B}_c$ to $\dot{B}_x$ is on the order of 35% under nominal operating conditions for this system. $\dot{B}_f$ contributes a correction on the order of a few percent; therefore, it is essential to include these two terms. Note from Eq. 10 that measurement of O_2 concentration in the exhaust gas is not required to estimate only the cell mass concentration. This measurement is necessary, however, if sugar and ethanol concentrations also have to be estimated.

An estimate of the specific growth rate $\hat{\mu}$ can be extracted from Eq. 5b by simply rearranging it to the form

$$\hat{\mu} = \frac{\hat{Y}}{\hat{x}} \tag{12}$$

Equations 6 and 10 represent the nonadaptive, nonlinear cell mass estimator. Koshy (1986) used a similar form for the cell mass estimation in a fed-batch fermentor with short fermentation times. This approach now will be extended significantly to improve the estimator's long-term robustness and flexibility to deal with different process operating conditions.

Design of a Multirate Adaptive Nonlinear Cell Mass Estimator

At this point the estimator still has an open-loop structure, and hence will lack robustness in the face of modeling errors. The biggest model uncertainty is expected to arise from the cell mass growth rate term $\hat{Y}$. Recalling the complexity of cell metabolism, it is questionable if the cell mass growth rate can be reliably estimated over very long periods of time, using two fixed parameters a and b, since several restrictive modeling assumptions have been made: The cell mass growth rate is modeled by a single reaction with constant stoichiometric coefficients; the cell mass composition is constant; and the CO_2 solubility and gas-liquid equilibrium constant are invariant. One can especially anticipate a slow composition change of the cells and the introduction of lag phases in the growth process as the process moves from one operating region to another. There are other factors that make adaptation highly preferable: The measured base addition rate in Eq. 10 is calculated based on the normality of the base solution and on the base pump calibration. Both these values can vary from run to run or even during a single run. However, if the parameters in Eq. 10 are adapted, then such errors can be completely absorbed within the parameter values.

It can be shown mathematically that the relative error of $\hat{Y}$ (w.r.t. its steady-state value) is the sum of all relative errors of parameters and measurements. Furthermore, an error e in $\hat{Y}$ is amplified by a factor of $1/D$, in its effect on the estimation offset. It is therefore crucial to keep model uncertainties and measurement noise levels to a minimum. Noise corruption can be reduced by filtering the measurement quantities; the effect of modeling error can be decreased through an adaptive scheme. A novel algorithm, where the estimation loop is closed with a proper feedback signal and where parameters are adapted, is now described.

The adaptation of the model parameters a and b requires a measurement of the cell mass growth, Y_m. Strictly speaking, this step is only possible at steady state, where the specific growth rate μ is equal to the dilution rate D. It then follows from Eqs. 3a and 5b that

$$Y_m = Dx_m \tag{13}$$

Measuring the cell mass concentration, x_m, at two or more steady states would then allow the parameters a and b to be computed. This method has, however, two serious disadvantages. First, it is seldom the case that fermentation systems achieve a precise steady state. Second, such a steady-state method would be very time consuming. Unfortunately, a dynamic measurement of the cell mass growth rate, Y_m, is not possible, because μ cannot be measured. An alternative approach is to use an estimate of the specific cell growth rate, $\hat{\mu}$ from Eq. 12, multiplying it by the measured cell mass concentration, x_m, to generate an expression analogous to a measurement of the cell growth rate

$$Y_m \approx \hat{\mu}x_m = a\left(\frac{\dot{B}_m}{V_L} - bC_gD - 8.26\cdot 10^{-4}D\right) + v(t) \tag{14}$$

where $v(t)$ is the error. Alternatively, x_m could be used in Eq. 12, but this approach would only amplify the off-line measurement error, leading to performance deterioration.

Consequently, the multirate adaptive nonlinear cell mass estimator consists of Eqs. 6, 10, 12 and 14. Figure 1 illustrates the feedback structure of the estimator and makes clear how

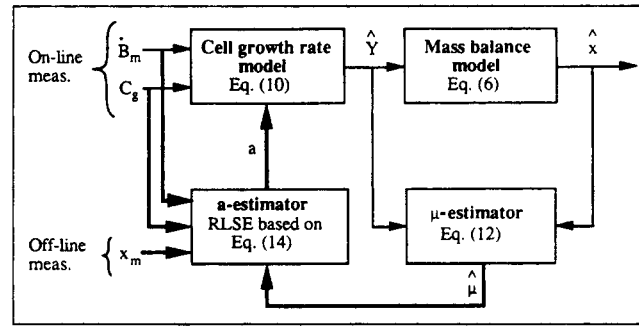


Figure 1. Feedback structure of adaptive nonlinear cell mass estimator.

—, frequent measurement signal paths; –, infrequent measurement signal paths.

new independent information is fed into the loop in order to eliminate possible bias in the estimated state variable. The bold signal paths signify infrequent information transmission, whereas all other paths denote that new information is transmitted frequently, every on-line sampling period Δt_{csl}. Off-line samples can be taken infrequently and even irregularly, because there is no specified sampling period embedded in Eq. 14. A recursive least-squares parameter estimator (RLSE) with constant forgetting factor using Eq. 14 is activated whenever an off-line measurement of cell mass concentration becomes available. It is, however, sufficient to update only the parameter a on-line; b is taken to be constant, an approach that will be justified below.

Simulation results using actual process data showed that even the imposition of constant parameters a and b produced a reasonable state estimator performance. Therefore, the necessary parameter adaptation can be expected to be small. Furthermore, a sensitivity analysis based on Eq. 14 with nominal process values typical for this fermentation system reveals that

$$\frac{\partial Y_m}{\partial(a/\bar{a})} \approx 2.0 \quad \text{and} \quad \frac{\partial Y_m}{\partial(b/\bar{b})} \approx 0.5 \tag{15}$$

These results imply that a relative error in the parameter b has one-fourth the effect on Y_m of the parameter a. It is therefore reasonable to lump all adaptation into a, leaving b constant. This approach, made possible through incorporation of *a priori* process knowledge, considerably improves the estimation quality in terms of higher convergence speed and lower computational load, while still guaranteeing good adaptation capability. A proper choice of the fixed parameter b is advisable in order to avoid excessive adaptation of the parameter a, which will cause deterioration in the estimation process. Initial a and b parameter values can be obtained from two or more steady-state data points. The adaptive estimator will subsequently compensate for modeling errors.

Analysis of the Recursive Least-Squares Algorithm

Since only parameter a in Eq. 14 is adapted on-line, the form of the recursive least-squares algorithm with exponential forgetting reduces to a very simple scalar form:

$$Y_m(t) = a(t)X(t) \tag{16}$$

$$P(t+1) = \frac{P(t)}{\lambda + P(t)X^2(t+1)} \tag{17}$$

$$a(t+1) = a(t) + P(t+1)X(t+1)(Y_m(t+1) - a(t)X(t+1)) \tag{18}$$

where X denotes the measurement function in Eq. 14 (function of base addition rate and CO_2 concentration); P denotes the covariance term (a scalar); λ is the forgetting factor. The product PX often is referred to as the Kalman gain. Such a scalar form is not only easy to code in a real-time estimation program, it has other advantages. The convergence is faster than in multiparameter estimators, and the parameter bias problem is reduced drastically with only a single parameter to estimate. The need for strong excitation is reduced to a minimum. Furthermore, analysis of such a least-squares estimator becomes very transparent because only scalar variables are involved.

There are several considerations that have to be incorporated in order to implement a recursive least-squares algorithm successfully. The algorithm requires initial values for the parameter estimate and for the covariance. It is obvious that the parameter convergence is fast when the initial guess is close to the correct value. In fact, dozens of samples may be required to commission the estimator if the initial guess is significantly in error. However, with initial values that are obtained by the steady-state method described above, the commissioning period is shortened considerably. Thus, it is advised to take off-line samples more frequently in the beginning, until reasonable cell mass estimates are obtained. Once convergence is achieved, the off-line sampling rate can be decreased.

Because of starting parameter uncertainty, the initial covariance value should be high to allow for fast convergence. It was found via simulation that too low an initial covariance value can be very detrimental to the convergence rate. On the other hand, no performance deterioration was observed when a very high initial value for the covariance was chosen. This characteristic can be explained by simple analysis: The very first update is only affected by $P(t)@t=1$, that is, $P(1)$, and not by $P(0)$. In addition, for $P(1)$ the following limit exists:

$$\lim_{P(0)\to\infty}\{P(1)\} = \lim_{P(0)\to\infty}\left\{\frac{P(0)}{\lambda + P(0)X^2(1)}\right\} = \frac{1}{X^2(1)} \tag{19}$$

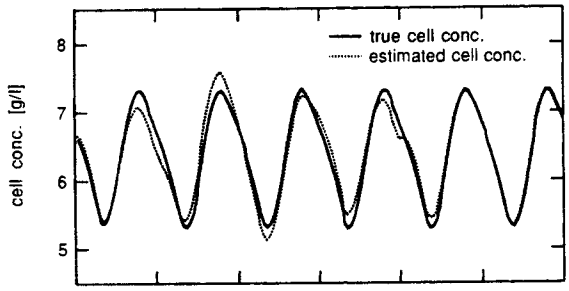
This limit can be calculated based on process knowledge; nominal parameter and measurement variable values can be used to evaluate the term in parentheses of Eq. 14 which represents X in Eq. 19. In this case, the limit of covariance $P(1)$ is suitable to give fast convergence. Therefore, $P(0)$ can be set arbitrarily high, as long as $P(1)$ is close to its limit.

Another important implementation consideration is covariance windup. The conditions that can lead to covariance windup are a forgetting factor less than unity and insufficient information, such as when operating at steady state. An analysis of this problem can be made easily by finding the limit of the covariance, assuming a constant measured value X:

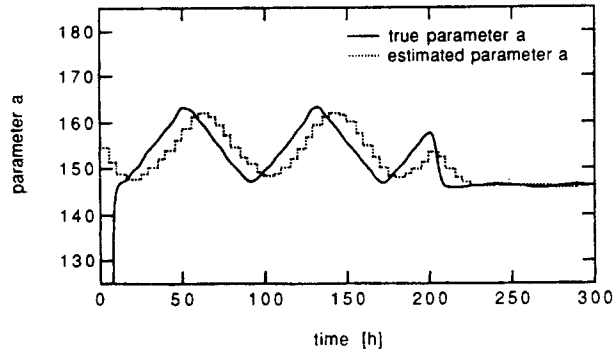
$$\lim_{t\to\infty}\{P(t)\} = \frac{1-\lambda}{X^2(\infty)} \tag{20}$$

Two features can be seen from inspection of Eq. 20. First, if the forgetting factor is unity, then the covariance tends to zero, and the estimator "falls asleep." Second, the covariance only tends to infinity if X is zero at steady state. This result could occur in the case when the estimator equations are expressed in deviation variables. In the present situation, absolute variables are used that never approach zero values. Therefore, the covariance has an upper bound even when no new information is arriving. Again, this limiting value can be calculated exactly from *a priori* process knowledge. For the fermentor system, the covariance limit at steady state is quite low, and there is no need to turn the estimator off for reasons of avoiding covariance blowup.

At this point the cell mass state estimator is fully ready for implementation and basically ready to be applied on-line. There are, however, two more issues that need to be checked, that

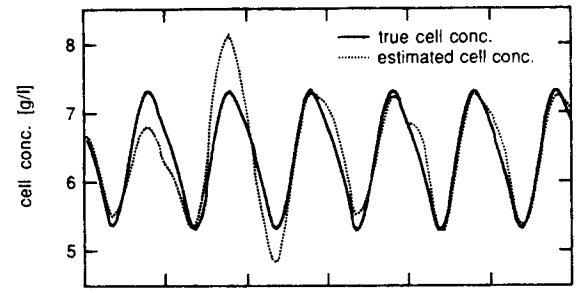


a. Cell concentration estimation error

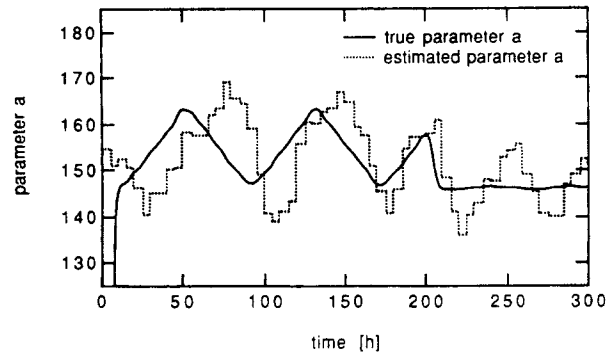


b. Parameter a estimation error

Figure 2. Unbiased cell mass estimation with no measurement noise.



a. Cell concentration estimation error



b. Parameter a estimation error

Figure 3. Unbiased cell mass estimation with measurement noise.

is, closed-loop stability and the capability to eliminate state estimation biases. Only simulations, with clearly defined conditions, can establish bias elimination and closed-loop stability.

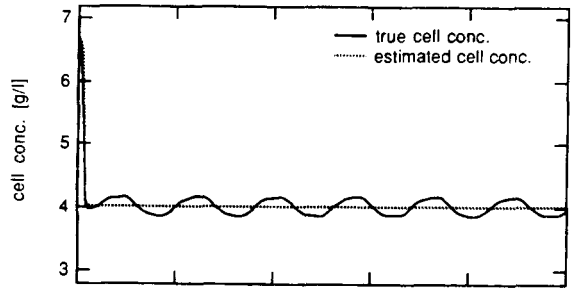
Simulation Results

A model/plant mismatch was first introduced by varying parameter a of the simulated process in an oscillatory fashion. Off-line samples were assumed to be available every 5 hours; optionally, they could be corrupted with Gaussian noise. By this means, it could be demonstrated that the adaptive feedback structure is capable of providing unbiased cell mass estimates. The open-loop process also was perturbed in this case by an oscillatory dilution rate. The forgetting factor was set to 0.6. Figure 2 shows, for the no-noise case, an unbiased tracking of the parameter a, which further yields an unbiased cell mass estimate. The parameter estimator does slightly lag behind the process, a direct consequence of the filtering characteristics of the recursive least-squares estimator. Decreasing the forgetting factor could decrease this lag, but it would also unreasonably amplify the measurement noise that is always present in a real process. This latter conjecture is validated by the responses in Figure 3 where the cell measurements are perturbed with noise [variance = 0.1 $(g/L)^2$]. After time 200 hours, when the process parameter a is set constant, all fluctuations in the estimate a are due to measurement noise, which is not attenuated enough by a forgetting factor of 0.6. For fairness, it should be noted that the measurement variance of 0.1 $(g/L)^2$ used in the simulations is much higher than expected with data from a real process. This choice was made to test the algorithm under extreme conditions.

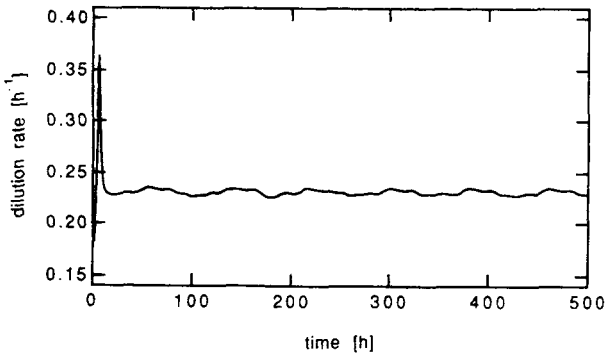
In a second set of simulation tests, the loop was closed using a nonlinear adaptive Generic Model Control (GMC) algorithm (Ramseier et al., 1991) to control the cell mass estimate. The setpoint response and the manipulated variable are shown in Figures 4a and 4b. The parameter a estimate constantly lags behind the corresponding value of the process parameter as shown in Figure 4d. Since a wrong estimate of parameter a implies a wrong state estimate, the same oscillations are seen in the true state variable. Because the state estimate is controlled and constant, and there is an oscillatory offset between real and estimated state variables, it is the real state that oscillates and not the estimated one. The amplitude of the state oscillation is a function of the parameter estimator lag and the strength of the control actions. The larger the estimator lag and the stronger the control actions, the higher the amplitude of the state variable response. In an extreme case, one can even see an oscillation in the controlled state variable estimate.

The important conclusion from these closed-loop simulations is that oscillations in the true cell mass are due to a parameter tracking lag; there is no indication of system instability for this process or of steady-state offset.

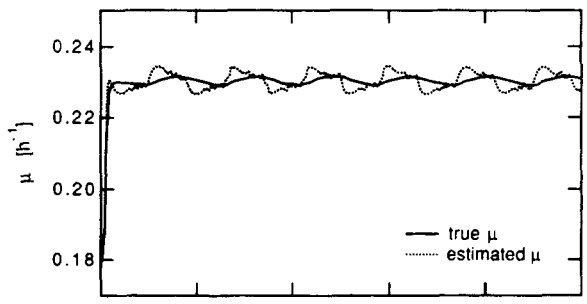
Another important design parameter is the off-line sampling period used for cell mass concentration measurements. Figure 5 shows how the off-line sampling period Δt_{ols} affects the estimator performance in the face of a process disturbance. This



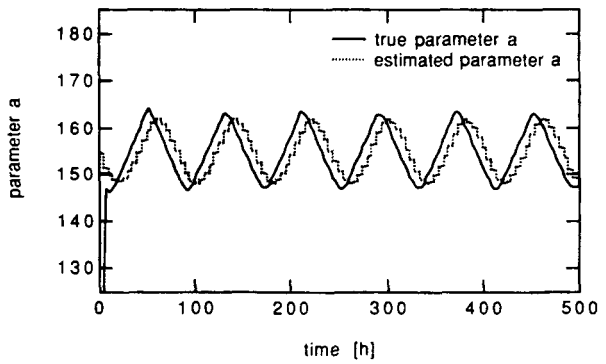
a. Output variable setpoint response



b. Response of manipulated variable

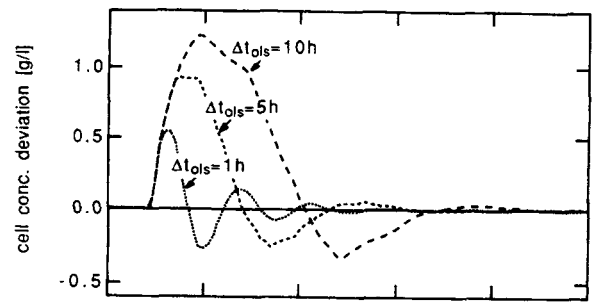


c. Parameter μ estimation error

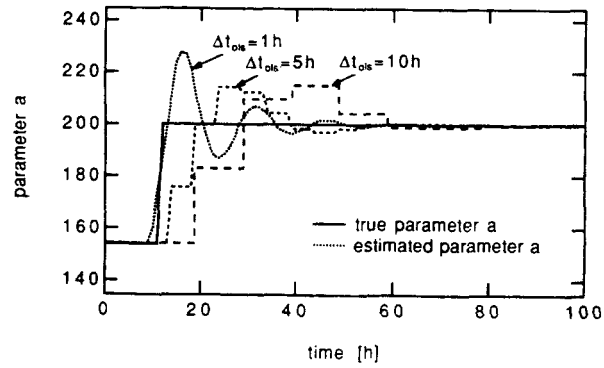


d. Parameter a estimation error

Figure 4. Closed-loop estimator performance for set point change.



a. Cell concentration estimation error



b. Parameter a estimation error

Figure 5. Performance deterioration for increasing off-line sampling period for a large modeling error.

situation was simulated by introducing a step change in the true process parameter a at time 10 hours. The estimation error converges to zero faster as the sampling period becomes shorter, but the response also becomes more oscillatory. It can also be seen from Figure 5 that the settling time is about 30 hours in the best case ($\Delta t_{ols} = 1$ h) using a forgetting factor of 0.6. Choosing a smaller forgetting factor could speed up the response, but it would also considerably amplify the measurement noise and therefore deteriorate the result. The above analysis indicates that the off-line sampling rate becomes a significant performance factor if drastic process changes must be accommodated.

Experimental Results

Off-line analysis of an open-loop step response. Off-line analysis of experimental data was used to obtain more insight into the physical process and to put the adaptive estimator in its final form for implementation. The dilution rate was changed in a stepwise fashion over a range of 0.2 to 0.3 h^{-1} to obtain open-loop step response data for base addition rate, CO_2 off-gas concentration, and off-line cell mass measurements. The base addition rate and CO_2 off-gas concentration, both signals with a considerable amount of variation/noise, had to be filtered. A conventional exponential filter was used, with filter time constants of 0.95 hours for base addition rate and 0.20 hours for CO_2 off-gas concentration. The high level of base

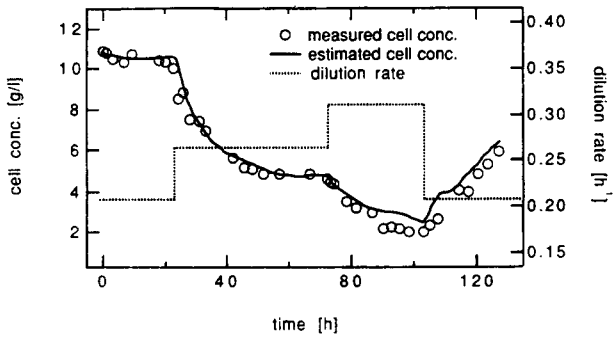


Figure 6. Cell mass estimator performance with fixed parameters.

($a = 180.6$, $b = 0.025$, std. dev. = 0.385 g/L).

addition rate filtering was necessary because of the discrete and on-off type pH control scheme.

Since actual flow rates could not be measured, the pump on-time was used to compute the base addition rate. An air bubble in the base feed tube could therefore lead to a temporary erroneous base addition rate reading. A spike filter was used to eliminate erroneous spikes in the base addition rate. The threshold was set to 28 mmol/h multiplied by the volume of the fermentor (in liters). This value was chosen well above the maximum attainable base addition rate due to cell growth, yet it was set low enough to eliminate the effect of even small bubbles.

Once all the process data were conditioned, they could be applied to the estimator equations, Eqs. 6, 10, 12, and 14. This approach allowed a reliable and quick off-line testing of the adaptive estimator. Design parameters could easily be changed to see the effect on the state estimator performance.

The limitation of a fixed parameter estimator can be easily seen in Figure 6 using *a posteriori* information. The fixed parameters a and b were obtained from the two steady states at 20 and 70 hours ($a = 180.6$ and $b = 0.025$). The estimator fits the measurements very well between these two steady states, but a clear offset can already be noticed outside this range. However, a fixed set of parameters a and b is seen to estimate the cell mass concentration reasonably over a period of 130 hours and over a variety of operating regions indicating that adaptation can be expected to be only moderate. Also, an off-line sampling period of about four hours, as was used in obtaining these results, seems to be appropriate.

To check the influence of the forgetting factor, a series of tests was conducted using off-line experimental data. The effect of a decreasing forgetting factor (increased forgetting) is demonstrated by comparing Figures 7 ($\lambda = 0.95$) and 8 ($\lambda = 0.10$). The effect can be seen at the sample points around 30 hours and around 45 hours. The lower the forgetting factor the better the local fit around 30 hours, but also the worse the prediction at 45 hours, after a long time span of no off-line samples. The estimator with a low forgetting factor thus gives too much weight to recent measurements, filtering less effectively as a consequence. In other words, it has the tendency to fit data locally very well but to yield bad predictions. These results make clear the trade-off between fast adaptation and noise sensitivity. Normally, this trade-off makes it hard to pick a value for the forgetting factor. We can resolve this problem, however, by incorporating *a priori* process experience. It can

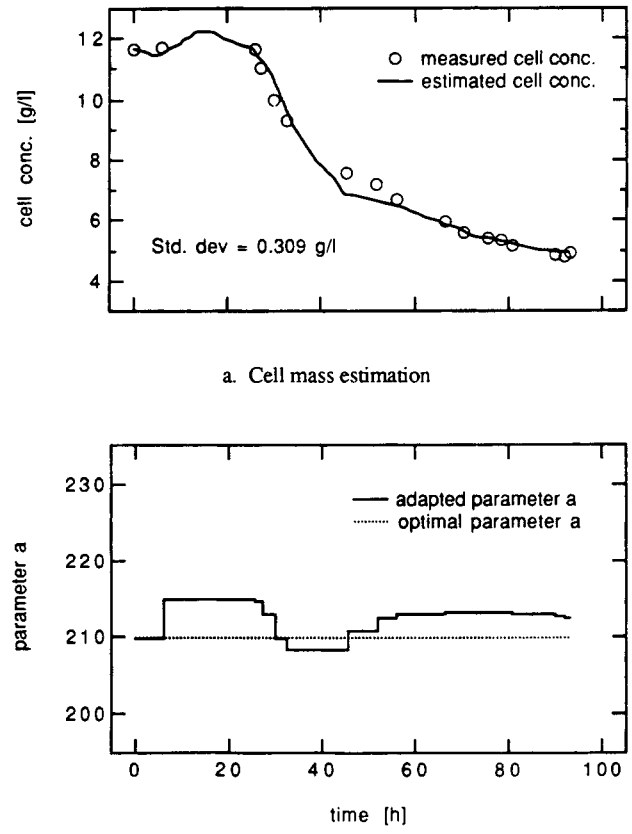
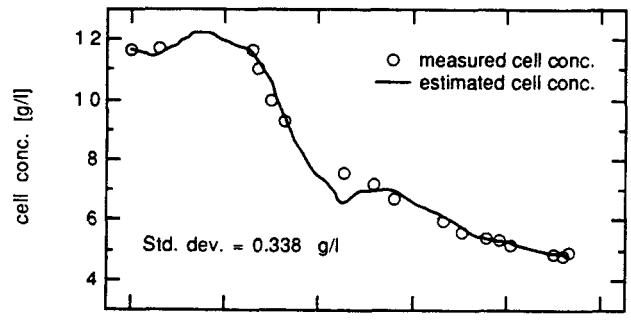


Figure 7. Adaptive cell mass estimator with large forgetting factor ($\lambda = 0.95$).

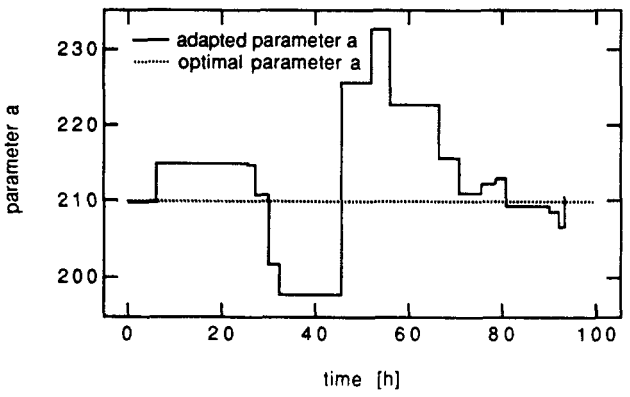
be conjectured from Figures 6 and 7 that the amount of expected adaptation will be small under normal operating conditions and it certainly can be stated that parameter changes requiring a low value of forgetting factor, for example, 0.1, are not expected. A much more reasonable choice would be a higher value, say 0.6, a compromise between high adaptation speed and low measurement noise amplification. It is, however, interesting to note that even low values of forgetting factor (even an unreasonable value of 0.1) do not lead to catastrophic estimation at all since the method is very insensitive to the choice of the forgetting factor, greatly facilitating implementation. The parameter estimates also are not biased and merely fluctuate around the value found to be optimal (for this particular response).

At this point it has to be stressed that additional process excitation, such as a Pseudo Random Binary Sequence (PRBS), is not necessary. The parameter a converges even when the estimator is only given data close to steady state.

Ydstie and Golden (1987) show that even a simple closed-loop system at steady state can exhibit slow parameter drifts which lead to bursts in the controlled output variable. Closed-loop simulations did not reveal such chaotic behavior. However, to be on the safe side in applications with the actual process, a simple means was provided to avoid such a phe-

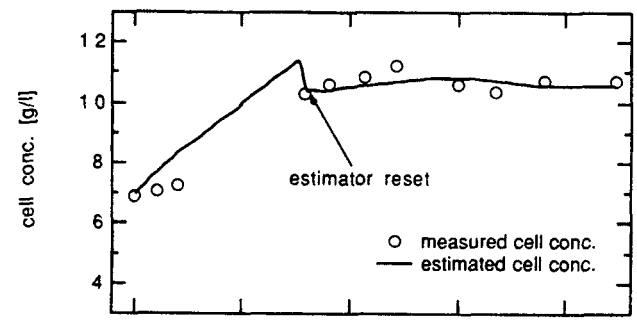


a. Cell mass estimation

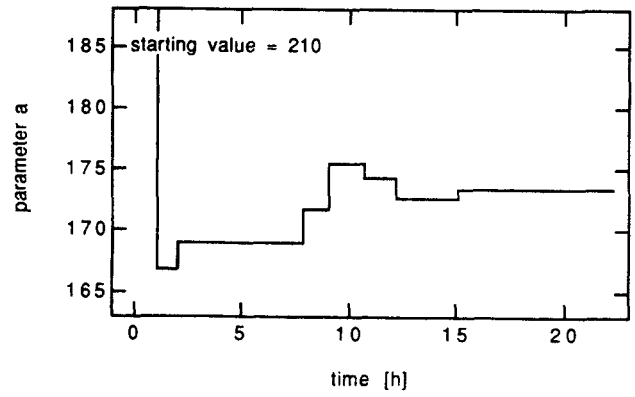


b. Adaptation of parameter a (b = 0.021)

Figure 8. Adaptive cell mass estimator with small forgetting factor (λ = 0.10).



a. Cell mass estimation



b. Adaptation of parameter a

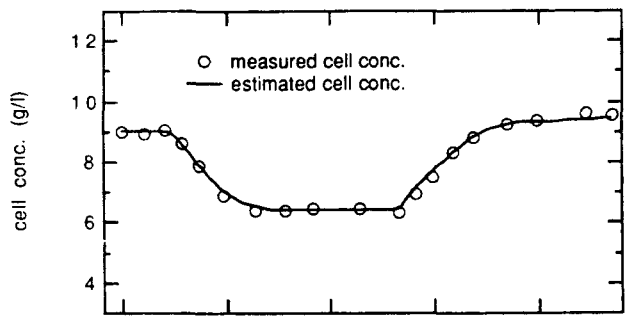
Figure 9. Typical commissioning period for the estimator under experimental conditions.

nomenon. It is common practice to define a prediction error band. If the prediction error falls into that so-called deadband, then the estimator is not activated. Prediction errors at the nominal steady state for the laboratory fermentor are of the order of ± 0.01 g/L·h and significantly higher in other regions. Therefore, the deadband was set to ± 0.02 g/L·h to add more security.

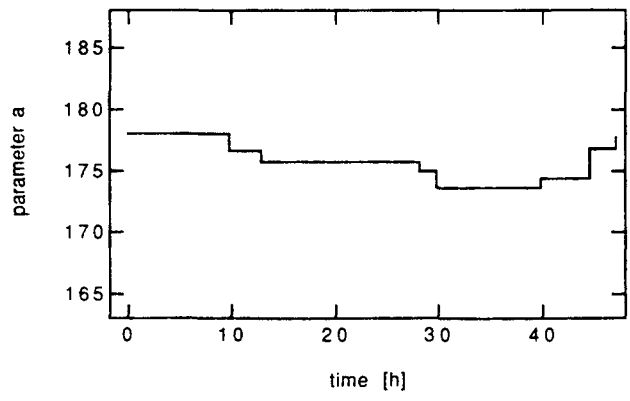
On-line results. The real check of the adaptive nonlinear cell mass estimator involves implementing it on-line with the actual yeast fermentation system. Figure 9 represents a typical commissioning pattern. Rapid off-line sampling is required for fast estimator convergence. A reasonable convergence is achieved after about one day. During this period it may be necessary to reset the cell mass estimate to the measured value, if the difference is too large. This technique can significantly enhance the convergence speed, as is shown in Figure 9. The forgetting factor for the commissioning period as well as for all subsequent experiments was set to 0.6. Also, the constant parameter b was set to 0.0213, a value found optimal from previous open-loop experiments.

Figure 10a shows the excellent estimator performance for two separate transient responses where the dynamics of the process are fast. Because the cell mass estimate is controlled in this case, large changes in the cell mass concentration lead to fast closed-loop responses. Even so, the estimator tracks the true cell mass concentration effectively. These results prove that the estimator can cope with fast process dynamics. It is also evident from the profile of parameter a in Figure 10b that the adaptation is quite small for this case. However, the profile of parameter a indicates clearly that a is smaller for low cell mass concentrations and larger for high cell mass concentrations. The irregular off-line sampling rate for the a-parameter adaptation is chosen with an average period of about three hours. At steady-state operation samples can be taken less frequently; during fast transients the sampling frequency has to be higher. To check the estimator performance, more cell mass samples were taken than actually used for the a-parameter adaptation. Figure 10c depicts the corresponding specific cell growth rate estimate of μ.

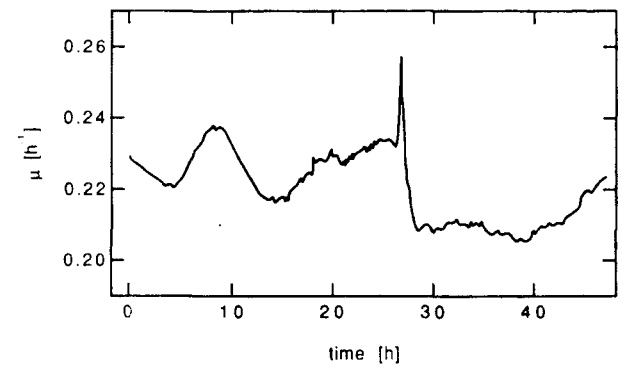
There are, however, situations where the estimator performance will be poorer. For example, the estimator performance is shown in Figure 11 for a sudden pH-step from 5.5 to 3.5. Such a large and sudden disturbance requires a large change in the parameter a in order to track the process. As previously shown in simulations (Figure 5), it takes a long time for the cell mass estimator to converge for such a large change in parameter a. It can be concluded that the relatively large cell mass estimator error is only due to lag in the a-parameter



a. Cell mass estimation

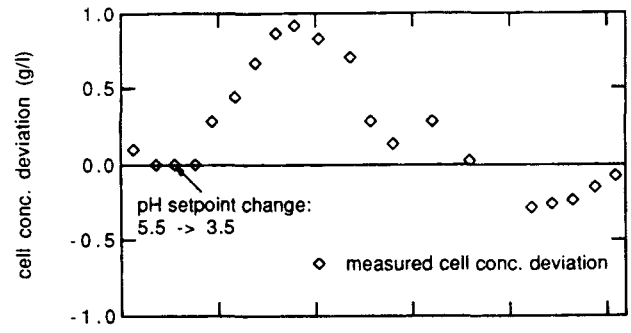


b. Adaptation of parameter a

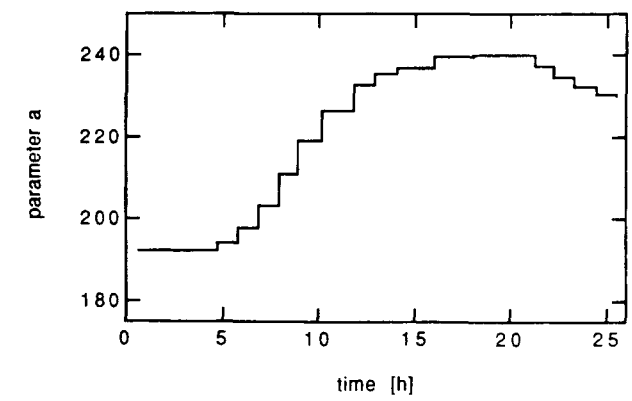


c. Estimation of parameter μ

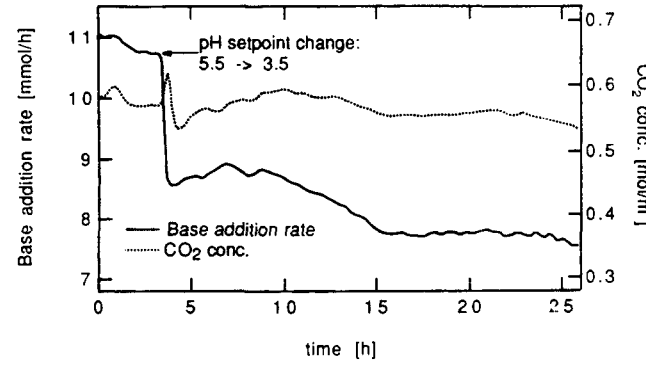
Figure 10. Cell mass estimator tracking under closed-loop (controlled) conditions.



a. Cell mass estimation error



b. Adaptation of parameter a



c. On-line measurements

Figure 11. Load change response characteristics of cell mass estimator (change in pH controller set point from 5.5 to 3.5).

estimation, which is a direct consequence of the filter characteristics of the recursive least-squares algorithm.

Similar behavior of the cell mass estimator was observed when a sudden downward dilution rate from 0.25 to 0 h^{-1} was introduced. Such a change is necessary to make a large setpoint change to higher cell mass concentration. The cells respond to such a drastic change in operating conditions by entering a "lag phase" during which they adapt to the new environment. Again, a large change in parameter a is required to adapt to this lag phase condition and the estimator performance becomes poor in the interim. It has to be noted, however, that such drastic (sudden and large) changes in setpoint or disturbances usually do not occur during normal fermentor operation.

Conclusions

The adaptive nonlinear cell mass estimator, as developed and rigorously tested here, exhibits good performance. It incorporates *a priori* process knowledge through its underlying nonlinear model. The parameter adaptation feature increases robustness, but can be effectively kept to a minimum. This approach significantly decreases the computational load and speeds up convergence compared to other published estimators. Also, the tuning effort is minimal. The multirate sampling structure provides a convenient way to feed back the best available (off-line) process measurement and provides both long-term robustness and the elimination of state variable offset at steady state. The limitations of this parameter estimator, shown experimentally, can be explained via theoretical analysis as arising from filtering constraints.

Structurally, the estimator incorporates elements from a number of existing methods. However, the favorable combination of a multirate adaptive technique and a nonlinear model with a single parameter to adapt provides a very practical estimation scheme. The successful incorporation of *a priori* process knowledge into the design is directly reflected in the very simple implementation and tuning requirements of this method and in its excellent operating characteristics under actual operating conditions.

Notation

a = cell mass estimator coefficient
b = cell mass estimator coefficient
B = base addition rate, mol/h
C_g = CO_2 concentration in gas phase, mol/m^3
D = dilution rate, h^{-1}
K = equilibrium constant
FW_x = formula weight of cell mass, g/mol
P = covariance matrix
Q_L = liquid flow rate, L/h
t = time, h
V_L = fermentor liquid volume, L
v = error
X = information vector for RLS algorithm
x = cell-mass concentration, g/L
Y = cell mass growth rate, g/L·h

Greek letters

$\gamma_1 \cdots \gamma_4$ = cell mass composition coefficients
Δt_{est} = on-line cell mass estimator sampling period, h
Δt_{ols} = off-line cell mass estimator sampling period, h
λ = forgetting factor
μ = specific cell growth rate, h^{-1}
$\nu_1 \cdots \nu_6$ = stoichiometric coefficients

Subscripts

g = gas phase
L = liquid phase
m = measured

Superscripts

$\hat{}$ = estimate
$\bar{}$ = steady-state value
$\cdot$ = time derivative

Literature Cited

Agarwal, M., and D. Bonvin, "Limitations of the Extended Kalman Filter for Batch Reactors," *Proc. IFAC DYCORD Symposium*, Maastrich, Netherlands (1989).

Agrawal, P., "An Experimental Study of Acid Production Rate Controlled Operations of a Continuous Fermentor," *Bioproc. Eng.*, **4**, 183 (1989).

Bellgardt, K., W. Kuhlmann, H. Meyer, K. Schügerl, and M. Thoma, "Application of an Extended Kalman Filter for State Estimation of a Yeast Fermentation," *Proc. IEE, Part D*, **133**, 226 (1986).

Chang, Y. K., and H. C. Lim, "Fast Inferential Adaptive Optimization of a Continuous Yeast Culture Based on Carbon Dioxide Evolution Rate," *Biotechnol. Bioeng.*, **35**, 8 (1990).

De Vallière, Ph., and D. Bonvin, "Application of Estimation Techniques to Batch Reactors: III. Modelling Refinements which Improve the Quality of State and Parameter Estimation," *Comput. Chem. Eng.*, **14**, 799 (1990).

Goodwin, G., and S. Sin, *Adaptive Filtering, Prediction and Control*, Prentice-Hall, 248 (1984).

Guilandoust, M., A. J. Morris, and M. Tham, "An Adaptive Estimation Algorithm for Inferential Control," *Ind. Eng. Chem. Res.*, **27**, 1658 (1988).

Halme, A., "Measurement and Estimation in Bioreactors," *One Day International Workshop on Control of Biotechnical Processes*, University of Newcastle upon Tyne, England (Oct., 1987).

Koshy, G., *On-Line Estimation and Optimization of Microbial Growth Based on Measurement of Acid Produced During Fermentations*, M.S. Thesis, University of California, Santa Barbara (1986).

Lewis, F., *Optimal Control*, John Wiley, New York (1986).

Ljung, L., "Asymptotic Behavior of the Extended Kalman Filter as a Parameter Estimator for Linear Systems," *IEEE Trans. Autom. Control.*, **24**, 36 (1979).

Ramirez, W., "Optimal State and Parameter Identification: An Application to Batch Fermentation," *Chem. Eng. Science*, **42**, 2749 (1987).

Ramseier, M., P. Agrawal, and D. Mellichamp, "Nonlinear Adaptive Control of Fermentation Processes Utilizing *a Priori* Modelling Knowledge," *J. of Proc. Contr.*, **2**, 129 (1992).

Stephanopoulos, G., and K. San, "Studies on On-Line Bioreactor Identification: I. Theory," *Biotechnol. Bioeng.*, **26**, 1176 (1984).

Tham, M. T., G. A. Montague, and A. J. Morris, "Application of On-Line Estimation Techniques to Fermentation Processes," *Proc. American Control Conference*, Atlanta, **2**, 1129 (1988).

Ydstie, B. E., and M. P. Golden, "Chaos and Strange Attractors in Adaptive Control Systems," *Proc. IFAC Congress*, Munich, Germany, **10**, 127 (1987).

Manuscript received Aug. 14, 1991, and revision received Aug. 24, 1992.