Advances in monitoring and state estimation of bioreactors

Ch. Venkateswarlu
Process Dynamics and Control Group, Chemical Engineering Sciences Division,
Indian Institute of Chemical Technology, Hyderabad 500 007

Received: 18 November 2003; accepted: 01 March 2004

The need to operate bioreactors in an efficient manner is becoming more relevant in today’s environment of changing process technology and increasing market competition. Better monitoring and control of bioreactors require reliable real-time available process variable information, while many of the important bioprocess variables cannot be measured on-line. On-line estimation of unknown bioprocess variables and incorporation of such variable information in process operation and control strategies can provide improved control performance with enhanced productivity. Bioprocess state estimation is thus an essential component to integrate with the process systems engineering tools such as, advanced process control, fault detection and diagnosis. This paper reviews the developments in state and parameters estimation of bioreactors from on-line monitoring and control perspective.

Keywords: Bioreactors, Bioprocesses, State estimation

IPC: Int.Cl: B 03 C 9/00, F 01 D 25/86

Introduction
Successful operation, control and optimization of biotechnical processes depend on reliable real-time available process variable information. In terms of control and optimization, bioprocesses are not much different than chemical processes. However, in bioprocesses, species are ill-defined and interact. The kinetics of bioreactions is not much understood yet. The process and disturbance dynamics are uncertain and vary during the course of operation. In spite of recent developments in optical, ion selective, and enzyme sensors, most of the variables in bioreactors cannot be measured reliably on-line. Process variables such as, temperature, pH, O₂ and CO₂ can be measured on-line. More informative measurements such as, measurements of metabolic products and intracellular metabolites that provide a more comprehensive description of the system cannot be easily measured on-line. The lack of on-line information concerning the process state impose limitations for effective control of bioreactors. Better monitoring and control of bioreactors require on-line estimation of process variables and parameters that can not be measured directly. The methods that provide on-line estimation of non-measurable process variables and parameters are called state estimators. The purpose of state estimation is to deliver reliable, real-time state variables, and parameters defining a process on the basis of available quantitative/qualitative knowledge of the process in conjunction with the known process measurements.

Bioprocess technology is currently employed for the production of several commodity and fine chemicals. Because of the complex nature of microorganisms growth and product formation in batch, semi-batch and continuous operations, the control of bioprocesses continues to present a challenging task. The on-line estimation of unknown bioprocess variables and their incorporation in the control law has significant advantages of compensating model mismatches and process uncertainties. This estimated state variable information can complement conventional sensor data/delayed measurement information, thus providing frequent feedback signals to controllers to achieve improved control performance. The estimator supported control schemes can lead to improved process operations with enhanced productivity. Bioprocess state estimation is thus an essential component to integrate with the process systems engineering tools such as, advanced process control, fault detection, and diagnosis. This paper presents an overview of various bioprocess state and parameter estimation methods, emphasizing recent advances from on-line monitoring.
and control perspective. The study is organized by classifying state and parameter estimation methods into different categories, which include methods based on balance equations, observers, Kalman filters, neural networks, and fuzzy reasoning.

**Balance Equation Based Methods**

By these methods, estimation is carried out using theoretically and experimentally derived relationships between the measured variables and the variables to be estimated. Simple calculations are used for estimation by neglecting measurement errors, modeling uncertainties, and instrument noise. Different nonlinear empirical/balance equation based methods have been derived for state and parameters estimations of biotechnical processes. Zabriskie and Humphrey\(^1\) have presented a method by which real-time estimates of biomass concentration and growth rate in fermentation processes are obtained by performing a material balance on oxygen and employing a kinetic model for molecular oxygen utilization. Mou and Cooney\(^2\) have described a computer aided methodology in which overall and instantaneous carbon balance equations are used to calculate the cell concentration and instantaneous specific growth rate in fed-batch penicillin fermentation. These carbon balancing equations together with a feedback control strategy are used to control the pre-selected production phase growth rate. San and Stephanoulos\(^3\) have presented relationships between the total rate of biomass growth and rate of ammonia addition to a fermentor for pH control. The derived relationships have been tested with non-biological acid-base continuous flow reaction systems and subsequently applied to the continuous yeast fermentation of glucose to ethanol. Grosz *et al.*\(^4\) have investigated the applicability of respiratory quotient measurement, a heat evolution measurement, and a commonly observed correlation between the respiratory quotient and product yield to on-line bioreactor identification and control. Experimental results and numerical simulations on the fermentations of yeast and *E.coli* support the theoretically derived rules. An on-line estimation technique for cell concentration of *Streptomyces avermitilis* fermentation, based on the on-line calculated oxygen uptake rate is presented by Gbewonyo *et al.*\(^5\). An on-line estimation method has been developed by Beluham *et al.*\(^6\) for biomass estimation, using maintenance equations. The biomass estimation data is used in a feedback control scheme of the baker’s yeast fermentation process.

**Observer Based Methods**

Observers are used as state estimators for deterministic linear systems. The knowledge of the mathematical model of the process along with the known measurements is used by the observers to estimate the process states. A block diagram, describing how the simulated or actual process measurements which can be used for the estimator is shown in Figure 1. Observers can be designed and implemented as linear observers or nonlinear observers. Most real systems often include system uncertainties arising due to time variation of parameters and nonlinearities. Linear observers cannot provide satisfactory performance for state estimation in nonlinear processes. Nonlinear observers, which exhibit robustness to systems uncertainties and adaptivity to system nonlinearities, are better suited for state estimation in nonlinear processes.

Various nonlinear observer based estimation schemes have been applied to predict the biological and physicochemical parameters needed for monitoring and control of biotechnical processes. The structure of the estimator is shown in Figure 2. Aborhey and Williamson\(^7\) have used asymptotic observers for state and parameter estimation of microbial growth processes. Dochain and Bastin\(^8\) have described the design of adaptive observers for state and parameter estimation of fermentation processes. The design of Luenberger observer is extended by Chu *et al.*\(^9\) to build a robust asymptotic nonlinear observer for state estimation of a biochemical reactor. Bastin and Dochain\(^10\) have discussed the advantages of observer based methods for estimation and control of bioreactors. Pomerleau and Perrier\(^11\) have presented an on-

![Fig. 1— Block diagram showing input to estimator](image-url)
line estimation technique using an observer for biomass concentration and an estimator for multiple specific growth rates of a baker’s yeast fed-batch process. Farza et al. have proposed nonlinear observer based estimators for on-line prediction of kinetic rates in fermentation processes. The performance of these estimators have been evaluated through numerical simulations and real-life experiments. Wang et al. have presented an adaptive observer for estimating the biomass concentration using the measurement of acid product in an anaerobic fermentation. The results of the observer are evaluated through experiments. Claes and Impe have employed an observer based estimator for the specific growth rate using viable on-line biomass measurements of fed-batch fermentation with bakers yeast in a stirred tank reactor. The design, tuning and implementation of the estimator have been illustrated by means of experiments. Bogaerts has proposed a hybrid estimation technique, based on exponential observer and asymptotic observer to estimate the state and identify the confidence of the kinetic model. Simulation study of a fed-batch bacterial culture is used to evaluate the performance of the estimator. Lubenova has presented adaptive observers for estimating time varying parameters like biomass growth rate and yield coefficient for oxygen consumption, and the state of biomass concentration in a class of aerobic bioprocesses using on-line measurements of oxygen uptake rate. The behavior of the estimators is studied through simulation.

Kalman Filter Based Methods

Kalman filter is used to provide optimal estimates of unmeasured and measured states for time varying linear systems in the presence of noise by combining information of a mathematical model of the process. The estimate obtained by Kalman filter at each time is the maximum likelihood estimate conditioned on all observations up to that time. The Kalman filter algorithm consists of two recursive steps. In the first step the process model is used to propagate the initial state estimates to that time at which the first measurement is available. In the second step the propagated model estimates are combined with the measurements to provide an updated or corrected estimates. Stephanopoulos and San have employed Kalman filter for on-line estimation of state variable of a biochemical reactor. Holmberg and Olsson have used Kalman filter for simultaneous on-line estimation of oxygen transfer rate and respiration rate using the measurements of dissolved oxygen concentration and air flow rate in an open aerator system. Howell and Sodipo have described a method in which both respiration rates and aeration coefficients can be determined on-line using only the measurement of dissolved oxygen concentration in an activated sludge aeration basin. Chattaway and Stephanopoulos have applied Kalman filter for state estimation of a recombinant cell culture. Ganovski et al. have discussed the implementation of a Kalman filtering procedure for state estimation of a batch Uriacase production process with Candida utilis by using unstructured model of the process.

Most biochemical systems of practical interest are inherently nonlinear and the use of a process model which reflects this essential nonlinear structure would prove to be more beneficial. As a consequence the estimation algorithms developed should either be inherently based upon the nonlinear structure or alternatively approximate the nonlinearities using a linear model with adaptation. Extended Kalman filter (EKF) is a standard nonlinear estimation technique, which has been used for combined estimation of states and parameters of nonlinear systems. By this method estimation is carried out by linearizing the nonlinear model equations around the current estimate and applying the Kalman filter to the linearized equations. Different versions of EKF are commonly used for state and parameter estimation in chemical processes. A block diagram of the estimation scheme shown Figure 3. Several applications of EKF have been reported for state and parameter estimation of biotechnical processes. Bellgradt et al. have employed EKF for state and parameter estimation of a yeast fermentation. Montague et al. have studied state estimation and parameter adaptive control in fedbatch fermentation for penicillin production. The EKF is used to estimate bio-
mass concentration in conjunction with the fermentation model by using carbon dioxide production rate and fermentor volume as measurements. Staniskis and Simutis have elaborated an adaptive measuring system for the main variables of a biochemical process using the theory of estimation. A recursive Kalman filter is used for state estimation and parameter identification based on a priori information of modeling and measurement noise and using the data of direct and indirect measurements. Ramirez has presented a general algorithm in which a sequential parameter identification algorithm is coupled with the Kalman filter state estimation algorithm. This algorithm has been applied for estimating the entire state of batch beer fermentation in the presence of uncertain model parameters. Pons et al. have applied EKF and a multivariable state estimator to the estimation of state variables of a biotechnological process in batch and fedbatch operation. The EKF has provided better global estimation of biomass, substrate and product. A two level extended Kalman filter has been applied by Venkateswarlu et al. for state and parameter estimation of batch beer fermentation. Sargantanis and Karim have used an iterative extended Kalman filter (IEKF) for estimating state variables like biomass content, dry matter content and moisture content using the measurements of total weight and carbon dioxide evolution rate. The state estimation scheme is used with an adaptive control structure for the control of solid substrate concentration. Myers et al. have applied EKF as a part of control system for simulation of a feedback biochemical reactor. The estimator utilizes noisy measurements of oxygen concentration and oxygen uptake rate with infrequent and time delayed measurements of biomass and substrate concentrations to estimate the state of the system. The state estimates are used with a feedback controller to track the optimal trajectory for biomass production. Gudi et al. have presented a multirate estimation algorithm for estimation of nutrient levels using frequent on-line measurements of carbon dioxide evolution rate, and off-line infrequent and delayed measurements of biomass and substrate concentrations. The estimation algorithm has been verified using simulations and industrial data from a fedbatch fermentation involving streptomyces species. The estimates are coupled with a nonlinear control law to track specified optimal nutrient trajectories.

Neural Network Based Methods

A neural network is a computing system made up of several simple, highly interconnected nodes, which process information by its dynamic state responses to external inputs. Neural network maps a set of input patterns onto a corresponding set of output patterns. The network accomplishes this mapping by first learning from a series of past examples defining sets of input-output correspondences for the given system. The network then applies what it has learned to new input patterns to predict the appropriate output. A block diagram of neural network estimator is shown in Figure 4. Inputs to the neural network estimator may consist of the manipulative inputs to the plant together with the other measured process variables. The corresponding process outputs provide the desired teacher signal, which trains the network. The difference between the desired output and the value predicted by the network is the prediction error. Itera-
tions are performed to minimize the prediction error such that the network is optimized. The trained network is then subjected to input patterns to predict outputs. Neural networks offer the opportunity to directly model the nonlinear processes, and estimate the values of relevant process variables.

In the recent past, there has been an increasing interest in the application of neural networks for state estimation and parameter identification of biotechnical processes. Karim and Rivera have presented a neural network methodology for the estimation and prediction of bioprocess variables using environmental and physiological information available from on-line sensors. Conjugate gradient method with unconstrained optimization is used to train the neural network. The authors found that the neural network estimator has shown better results for bioprocess state estimation. Thibault et al. have presented neural network computational algorithms for the dynamic neural modeling of bioprocesses. The dynamic neural model is used for the prediction of key bioprocess variables. Their simulation results of a continuous stirred tank fermentor have shown the distinctive ability of neural networks over traditional methods. Breusegem et al. have presented neural network techniques based on sliding window learning schemes for on-line prediction of fermentation variables. I et al. have applied a moving window neural network for dynamic modeling and on-line estimation of consumed sugar, cellmass, and product concentration in L-lysine fed batch culture. Teisser et al. have used several models built with static and dynamic neural network configurations with the objective of real time estimation and prediction of yeast concentration during growth. Syu and Hou have proposed neural networks with dynamic learning and prediction for the identification of a fermentation system producing mainly 2,3-butanediol. Shene et al. have presented neural network based estimator for the prediction of the state of Zymomonas mobilis CP4 fermentations.

**Fuzzy Reasoning Based Methods**

Fuzzy set theory is a mathematical tool for treating uncertain, semi-qualitative and linguistic information. The concept of graded membership of fuzzy set theory provides a basis of a systematic way of describing and manipulating the imprecision and vagueness of the natural phenomena. Rule based fuzzy reasoning is one of the commonly used approach of the fuzzy set theory. In this approach, the input-output mapping of a system is determined by a collection of fuzzy If-Then rules and by a corresponding inference mechanism. Fuzzy logic based approach has greater flexibility to capture the various aspects of incompleteness or imperfection about a system.

Fuzzy state estimators have been developed for estimating the unknown bioprocess states and parameters. Dohnal has used a fuzzy modeling approach to predict the specific yield of a fermentor using dilution rate and a growth rate as known variables. Postlewaite has presented a fuzzy state estimator for predicting specific growth rate of a fed-batch bakers yeast fermentation using substrate, inhibitor and ethanol concentrations, and dilution rate as measurements. Satish et al. have applied a fuzzy estimator for determining the concentration of plasmid bearing microorganisms in a continuous bioreactor.

**Other Methods**

Ko et al. have presented a least square procedure for on-line estimation of mass transfer coefficient and oxygen uptake rate and used them with an adaptive controller designed to control the dissolved oxygen concentration in the aerator of a wastewater treatment plant. Dochain and Bastin have given a nonlinear adaptive control approach for bacterial growth systems. The authors observed that the identified growth rate and yield coefficient are useful for monitoring the state of the biomass. Dochain and Pauss have used an estimation algorithm for the specific growth rate in an ethanol fermentation process. An on-line gas analysis based approach is illustrated by Petkov and Davis for estimating biomass during batch fermentation. Tenno and Uronen have formulated the problem of state estimation for a large scale wastewater treatment system using a distributed parameter stochastic model. Gomersall et al. have used a cluster analysis based approach for supervision and on-line state estimation of an industrial antibiotic fermentation process. Yuan et al. have described an algorithm for on-line estimation of maximum specific growth rate of an autotrophic biomass based on a general nitrification process model and applied to activated sludge systems. Differential evolution and genetic algorithms are optimization routines that operate in a similar manner to natural genetic selection. Na and coworkers have employed genetic algorithms to optimize and control the fed-batch growth of *S.cerevisiae* on glucose and estimate the reactor parameters on-line. Chou and Wang have used a hy-
brid differential evolution algorithm as an approach to state estimation in a system involving E.coli fermentation for the production of recombinant protein.

Conclusions

The study focuses on reviewing a wide range of approaches, including recent knowledge based techniques for state and parameter estimation in biotechnical processes. Balance equation based methods cannot provide reliable estimates in the presence of unpredictable changes in process parameters and noise in the measurements. Observability plays a key role for state and parameter estimation by model based estimators such as Kalman filters and observers. The existence of Observability in a qualitative sense indicates that measurable outputs contain useful information of all the state variables. Since most of the biotechnical processes are nonlinear, nonlinear estimation methods such as, nonlinear observers and extended Kalman filters are found useful for state and parameter estimation. Nonlinear observers and extended Kalman filters employ process models which reflect the essential nonlinear structure in a nonlinear or linearized fashion. Inspite of its apparent success, care should be taken in designing extended Kalman filter since it involves assumptions like the use of linearized process model and requires the knowledge of noise covariance matrices. Nonlinear observers can handle process nonlinearities due to involvement of nonlinear model structure in their design. However, care must be taken in designing nonlinear observers in the presence of unknown process parameters.

In certain situations, developing and validating a mathematical model is difficult and time consuming for bioprocess that have complex dynamics with unpredictable changes in process parameters. To overcome such difficulties, knowledge based tools such as, neural networks, fuzzy logic and genetic algorithm have emerged as alternate estimating methods. Neural networks have the advantage of adaptive learning behavior with the ability to automatically determine the relationships between the estimation variables and the estimator input variables. Neural networks have limitations as they require large amount of training data and provide poor estimates for the input data falling outside the boundaries defined by training examples. Fuzzy logic provides flexible means of building nonlinear relationships for the process variables, especially for complex nonlinear processes. But fuzzy reasoning requires prior knowledge of known behaviors between the estimation variables and the estimator input variables. Genetic and evolutionary algorithms also show much potential for bioreactor state estimation. The advantages and disadvantages of different state estimation methods are briefly presented in Table 1.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Supporting process knowledge</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance equation based methods</td>
<td>Input-output relations in the form of empirical equations</td>
<td>Involve simple calculations based on approximate models</td>
<td>Can not provide reliable estimates in the presence of process uncertainties and measurement noise</td>
</tr>
<tr>
<td>Observer based methods</td>
<td>Mathematical model</td>
<td>Provide state estimation in deterministic systems. Nonlinear observers effectively deal with nonlinear processes</td>
<td>Difficult to incorporate stochastic disturbances. Accurate model development is a major task</td>
</tr>
<tr>
<td>Kalman filter based methods</td>
<td>Mathematical model</td>
<td>Provide accurate estimation of states and parameters. Deal with stochastic disturbances and measurement noise</td>
<td>Developing and validating rigorous process models require much time and effort</td>
</tr>
<tr>
<td>Neural network based methods</td>
<td>Heuristic data in the form of empirical nonlinear correlations</td>
<td>Provide better performance with noisy and in-complete data. Neural networks are more adaptive</td>
<td>Poor generalization capability outside the training range. Require more training data</td>
</tr>
<tr>
<td>Fuzzy reasoning based methods</td>
<td>Heuristic knowledge in the form of production rules</td>
<td>Useful alternatives for processes which are complex, imprecise and vague</td>
<td>Require good understanding of a process to setup a complete and consistent rule base</td>
</tr>
</tbody>
</table>
So far, more efforts have been devoted towards theoretical research for developing estimation algorithms. Although some of the estimators are tested experimentally, more attention is required for their real time implementation together with model based controllers. Further efforts are also required towards theoretical research and practical implementation of knowledge based approaches in order to overcome the mathematical modeling difficulties associated with complex bioprocesses. While the study of state estimation to small scale systems is more common, its application to large scale processes requires more attention. The advancement in state and parameter estimation methods leads to better monitoring and control of bioreactors as it is an essential component in advanced control, gross error detection, and process fault diagnosis.

References