SRC-TR-87-70

CHEMICAL PROCESS SYSTEMS LABORATORY

An Expert Approach Towards State Estimation of Bioreactors

by

P. A. Minderman, Jr.
T. J. McAvoy
L. L. Gasner

RESEARCH REPORT

CHEMICAL PROCESS SYSTEMS ENGINEERING LABORATORY

AN EXPERT APPROACH TOWARDS STATE ESTIMATION OF BIOREACTORS

P.A. Minderman, Jr.
T.J. McAvoy
L.L. Gasner

A CONSTITUENT LABORATORY OF
THE SYSTEMS RESEARCH CENTER

THE UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742

AN EXPERT APPROACH TOWARDS STATE ESTIMATION OF BIOREACTORS

P. A. Minderman, Jr.
T. J. McAvoy
L. L. Gasner

Department of Chemical and Nuclear Engineering
University of Maryland
College Park, Maryland 20742

For presentation at the 1986 Annual Meeting of the
American Institute of Chemical Engineers

Miami Beach, Florida
November 2-7, 1986

Abstract

Previous investigators have demonstrated the theory and applications of macroscopic elemental and material balances to the estimation of bioreactor states. However most previous applications have been limited to processes which have only one growth phase. An expert systems knowledge base is presented which is used to extend this empirical estimation technique to processes which have multiple phases (e.g. a lag phase, a growth phase, and a production phase). This approach is illustrated by a simple example from the literature.

Introduction

The benefits of monitoring process variables have been well documented, and these benefits are realized because of quality control, process control, and/or research and development projects. Industrially this monitoring is generally implemented by measuring the desired process variables. For bioreactors these measurements are often expensive and time consuming. Thus there are incentives to estimate some of these variables; these incentives are to provide a control system with more timely information and to lower analysis costs while maintaining desired quality levels.

In the area of on-line estimation of bioreactors, many previous investigations can be classified into three groupings:

- state and parameter estimation based on detailed structured models (e.g. Svrcek *et al.* (1974), Staniskis and Levisauskas (1984))
- parameter estimation based on discrete input-output models (e.g. Lundell (1982), Rolf and Lim (1985), Harmon *et al.* (1984))
- empirical state estimation based on simple models (e.g. Cooney *et al.* (1977), Wang *et al.* (1977), Mou and Cooney (1983), Stephanopoulos and San (1984))

Because of a lack of reliable general structured models and of the large measurement requirements to estimate the numerous variables, the state and parameter estimation problem will not be addressed. Furthermore because the discrete input-output models are limited to process control systems, they will not be considered here. The authors feel that the simpler models yield the greatest information relative to the effort expended therefore have the greatest

potential for industrial control applications.

Background

Most previous empirical state estimation has been based on elemental balances and quasi-steady state gas material balances. For example, one postulates a chemical reaction of the form:

$$a\ carbonsource + b\ O_2 + c\ nitrogensource \rightarrow biomass + e\ H_2O + f\ CO_2$$

and then solves for these stoichiometric coefficients on-line by using the component empirical formulas, the four elemental balances, and the measurements of the carbon dioxide evolution rate and the oxygen uptake rate. This approach has been demonstrated by Wang *et al.* (1977) for the production of bakers yeast, by Ziegler and Humphrey (1979) for the production of single cell protein from ethanol, and by Stephanopoulos and San (1984) for the fed-batch production of ethanol from yeast. In each of these examples, the estimates agreed well with analytical measurements of biomass, substrate, and product (when applicable) concentrations during a specified phase (e.g. exponential growth) of the fermentation. Herein lies one problem limiting the extension of this approach to processes that have multiple phases (e.g. a lag phase, a growth phase, and a product phase). It is proposed that an expert systems approach can resolve this problem.

Expert's Knowledge Base

Experts have analyzed and modeled entire microbial processes; however, the success in modeling a complete general microbial process has been quite limited. More often experts resort to analyzing specific phases of a particular fermentation. Over the years, a number of serial phases have been identified, for example:

- lag phase upon innoculation
- lag phase between growth on multiple substrates
- lag phase between growth phase and production phase
- growth phase on a simple single substrate

- growth phase on a single inhibiting substrate
- growth phase on a single substrate and a competitive inhibitor
- growth phase on a single substrate and an uncompetitive inhibitor
- growth phase on a single substrate and a noncompetitive inhibitor
- stationary phase
- death phase

Other phases have been classified which may be in parallel with a growth phase:

- production phase of a growth associated product
- production phase of a nongrowth associated product
- production phase of an inducible product
- production phase of a product subject to feedback inhibition
- production phase of a product subject to feedback repression
- production phase of a product subject to catabolite inhibition
- production phase of a product subject to catabolite repression

Traditionally experts have analyzed microbial processes by considering which of the phases apply to their particular application. Consequently an expert approach is proposed in which one identifies the phases that apply to a given process, and applies previously published empirical estimation procedures to each phase independently. Unfortunately this implies that every different fermentation may require a unique set of estimators; however, because the bulk of the development effort is in the development of an estimator and because current software allows easy editing of production rules, this is seen as a minor inconvenience.

As this methodology is an on-line procedure, one must address the question of when to switch from one phase to another. Again the experts have identified this switch by some relatively easy measurements, for example:

- significant change in respiratory quotient, RQ
- significant change in carbon dioxide evolution rate, R_{CO_2}
- significant change in dissolved oxygen
- significant change in pH

- addition of an inducer

Unfortunately this measurement set is not nearly as exhaustive as the set of possible phases, hence more expensive and time consuming off-line measurements and/or estimated measurements may also be required to identify the phase switches. A simple example is used to illustrate this approach.

Example : Batch Growth of *Saccharomyces cerevisiae* on Glucose

Consider the batch growth of *Saccharomyces cerevisiae* on glucose as presented by Mor (1969) and shown in Figure 1. One can see that there are four distinct phases during this batch:

1. lag phase ($0 \leq t < 2$ hours)
2. growth on glucose (glycolysis) ($2 \leq t < 10$ hours)
3. growth on ethanol (respiration) ($10 \leq t < 19$ hours)
4. stationary phase ($19 \leq t$ hours)

Blindly applying the previous empirical approach yields the following chemical reaction:

$$a\ C_6H_{12}O_6 + d\ C_2H_5OH + b\ O_2 + c\ NH_3 \rightarrow$$

$$CH_{1.666}O_{0.51}N_{0.168} + e\ H_2O + f\ CO_2 + g\ C_2H_5OH$$

However these coefficients cannot be identified because the system of equations is singular (d and g are not independent). Previous investigators have solved this problem by considering only the glycolysis phase for which d equals zero.

The phases which apply to this example can be verbalized as a set of production rules as follows:

If (lag phase) then

do not estimate states - assume states are constant

Else if (glycolysis) then

estimate states based on the reaction:

$$a\ C_6H_{12}O_6\ +\ b\ O_2\ +\ c\ NH_3\ \rightarrow\ CH_{1.666}O_{0.51}N_{0.168}\ +\ e\ H_2O\ +\ f\ CO_2\ +$$

$$g\ C_2H_5OH$$

Else if (respiration) then

estimate states based on the reaction:

$$a'\ C_2H_5OH\ +\ b'\ O_2\ +\ c'\ NH_3\ \rightarrow\ CH_{1.666}O_{0.51}N_{0.168}\ +\ e'\ H_2O\ +\ f'\ CO_2$$

Else if (stationary phase) then

stop estimation - assume states are constant

End

An important assumption is that the empirical biomass composition is independent of the substrate. Stephanopoulos and San (1984) demonstrated that this composition is a weak function of specific growth rate for *S. cerevisiae*, so the assumption appears plausible.

Fortunately, off-gas composition measurements, Figure 2, allow one to discriminate between each of these phases. However the use of off-gas analysis to discriminate between respiration phase and stationary phase is difficult as the measurement noise becomes quite large. Better results are obtained when one uses an estimated measurement for this discrimination. Thus the production rules for this example become:

If (R_{CO_2} is decreasing AND the process is starting) then

lag phase

Else if (RQ is "high") then

glycolysis phase

Else if (RQ is "low") then

respiration phase

Else if (estimated [ethanol] = 0) then

stationary phase

End

Mathematical Formulation of *S. cerevisiae* Estimation Problem

Glycolysis Phase

Measurement Development

The basis of the empirical approach is a chemical reaction:

$$a\ CH_2O\ +\ b\ O_2\ +\ c\ NH_3\ \rightarrow CH_{1.666}O_{0.51}N_{0.168}\ +\ e\ H_2O\ +\ f\ CO_2\ +\ g\ C_2H_5OH$$

and the associated elemental balances:

$$\begin{aligned} a &= 1 + f + 2g \\ 2a + 3c &= 1.666 + 2e + 6g \\ a + 2b &= 0.51 + e + 2f + g \\ c &= 0.168 \end{aligned}$$

Further one can use the following quasi-steady-state balances:

$$R_x = \frac{1}{b} R_{O_2} \frac{M_{O_2}}{1000} = \frac{1}{f} R_{CO_2} \frac{M_{CO_2}}{1000} = \frac{1}{c} R_{NH_3}$$

$$RQ = \frac{R_{CO_2}}{R_{O_2}}$$

Combining these six equations yields the following on-line solution for the stoichiometric coefficients:

$$\begin{aligned} b &= 0.168 \frac{R_{O_2}}{R_{NH_3}} \frac{M_{O_2}}{1000} \\ f &= 0.168 \frac{R_{CO_2}}{R_{NH_3}} \frac{M_{CO_2}}{1000} \\ g &= f - b - 0.0355 \\ c &= 0.168 \\ a &= 1 + f + 2g \\ e &= 0.419 + f - g \end{aligned}$$

Thus the measurements for the estimator become:

$$\begin{aligned} R_x &= \frac{1}{b} R_{O_2} \frac{M_{O_2}}{1000} \\ Y_{gluc} &= \frac{0.8061}{a} \\ Y_{EtOH} &= \frac{0.5254}{g} \end{aligned}$$

Process Model

The process model is assumed to follow the presentation of Stephanopoulos and San (1984):

$$\begin{aligned}
\dot{x} &= \mu\, x \\
\dot{s} &= \frac{-1}{Y_{gluc}}\, \mu\, x \\
\dot{p} &= \frac{1}{Y_{EtOH}}\, \mu\, x \\
\dot{\mu} &= C + w_1 \\
\dot{C} &= w_1 \\
\dot{Y}_{gluc} &= w_2 \\
\dot{Y}_{EtOH} &= w_3
\end{aligned}$$

Respiration Phase

In a similar fashion, the equations for respiration phase follow the development of those for glycolysis phase.

Measurement Development

Chemical Reaction:

$$a'\ C_2H_5OH + b'\ O_2 + c'\ NH_3 \rightarrow CH_{1.666}O_{0.51}N_{0.168} + e'\ H_2O + f'\ CO_2$$

Elemental Balances:

$$\begin{aligned}
2a' &= 1 + f' \\
6a' + 3c' &= 1.666 + 2e' \\
a' + 2b' &= 0.51 + e' + 2f' \\
c' &= 0.168
\end{aligned}$$

Gas material balances:

$$R_x = \frac{1}{b'} R_{O_2} \frac{M_{O_2}}{1000} = \frac{1}{f'} R_{CO_2} \frac{M_{CO_2}}{1000}$$

$$RQ = \frac{R_{CO_2}}{R_{O_2}}$$

Combining these five equations yields the following on-line solution for the stoichiometric coefficients:

$$b' = \frac{-0.5355}{1 - RQ}$$

$$f' = b' RQ$$

$$c' = 0.168$$

$$a' = 1 - f'$$

$$e' = 0.919 - 0.5 f$$

Thus the measurements for the estimator become:

$$R_x = \frac{1}{b'} R_{O_2} \frac{M_{O_2}}{1000}$$

$$Y_{EtOH} = \frac{0.5254}{a'}$$

Process Model

$$\dot{x} = \mu x$$

$$\dot{s}' = \frac{-1}{Y_{EtOH}} \mu x$$

$$\dot{\mu} = C + w_1$$

$$\dot{C} = w_1$$

$$\dot{Y}_{EtOH} = w_2$$

Optimal Estimation Procedure

Each set of states are estimated with a discrete extended Kalman filter (Jazwinski, 1970) as discussed by Stephanopoulos and San (1984); time varying forgetting factors (Ydstie, 1985) are included also. This estimator is not sensitive to reasonable deviations in the initial conditions or to reasonable white noise added to the estimator measurements.

Results of Estimation

The results of the estimation are compared with literature data in Figure 3, and the results are quite satisfactory. The statistics of this fit follow.

Glycolysis state residuals:

biomass	mean = -0.057 g/l	σ = 0.034 g/l
glucose	mean = -0.181 g/l	σ = 0.128 g/l

ethanol mean = 0.082 g/l σ = 0.094 g/l

Respiration state residuals:

biomass mean = -0.095 g/l σ = 0.280 g/l

ethanol mean = -0.004 g/l σ = 0.221 g/l

These residual means are well within experimental accuracy; for example, the glucose estimate is the least accurate because it is being compared with the least accurate experiment data.

By closely examining Figure 3, one can see a feature that may bother some, namely that the estimates are discontinuous. Admittedly this is not reasonable for many microbial processes; however, it seems like a reasonable price to pay for the increased flexibility allowed by this approach. Nevertheless, one must be careful when studying states that are not being updated (*e.g.* biomass in stationary phase and glucose in respiration phase); the state estimates appear to be biased when in fact the estimates are being linearly extrapolated. As in any situation, one must be careful with extrapolations.

Conclusions

An expert systems approach is presented which allows the extension of the principle of estimating bioreactor process variables through the use of macroscopic elemental and material balances to multi-phase processes. This approach is illustrated through a simple literature example. Although the limitations of estimating based on simple models remain in that only a relatively small number of states can be estimated without significantly increasing the measurement requirements, the potential applications of this approach are numerous as many industrial processes are multi-phase. The relative simplicity, process flexibility, and ease of extension of this approach continue to make simple estimation models promising for industrial control applications.

Figure 1 – *Saccharomyces cerevisiae* from glucose (Mor, 1969)

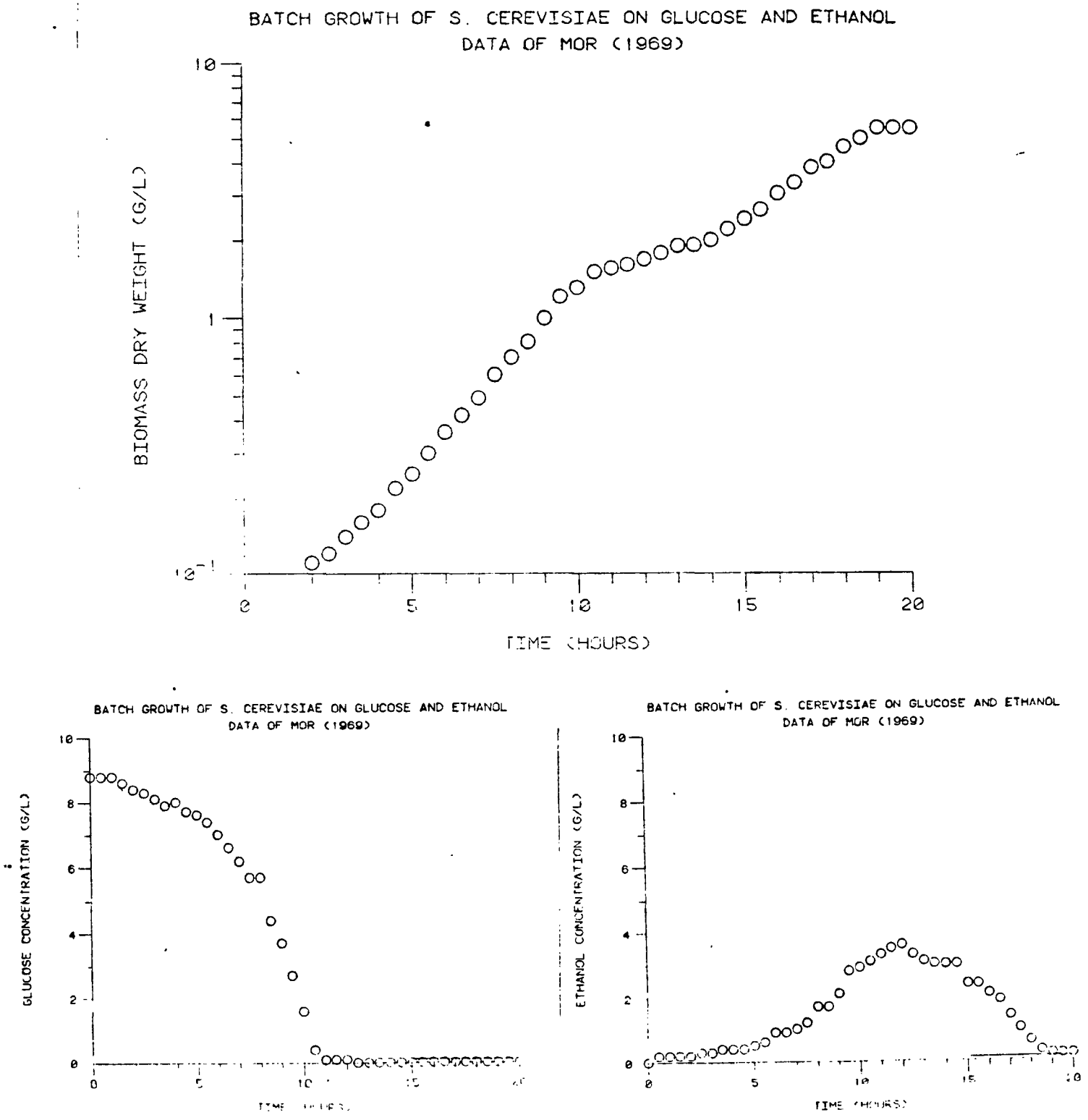
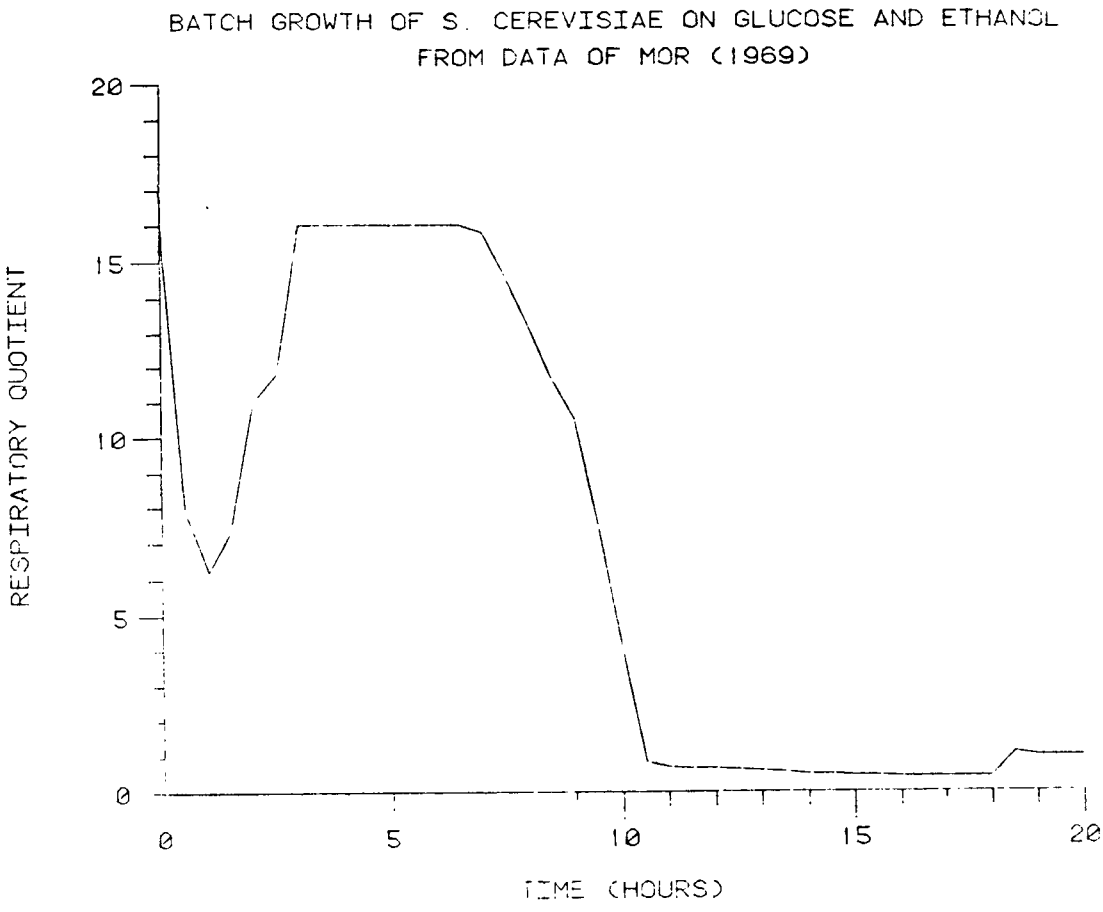


Figure 2 – Gateway Measurements for On-Line Estimation



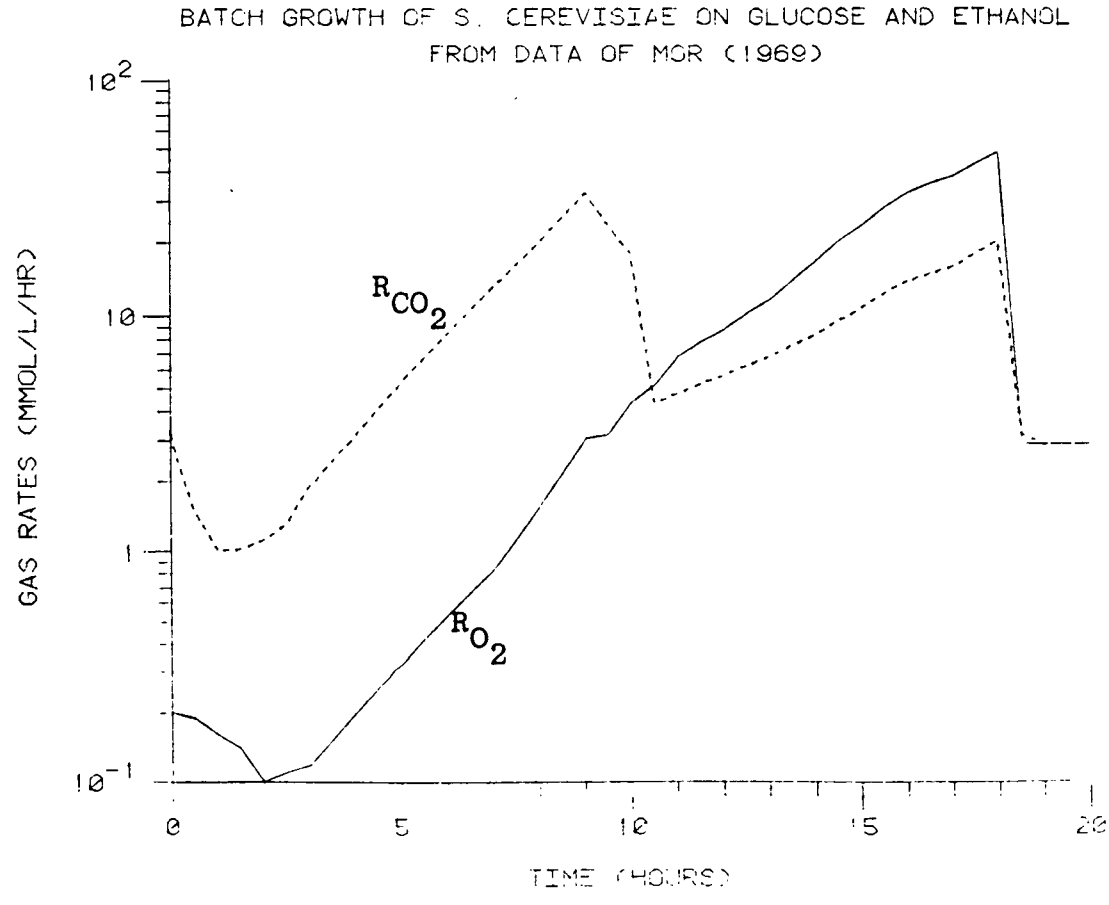
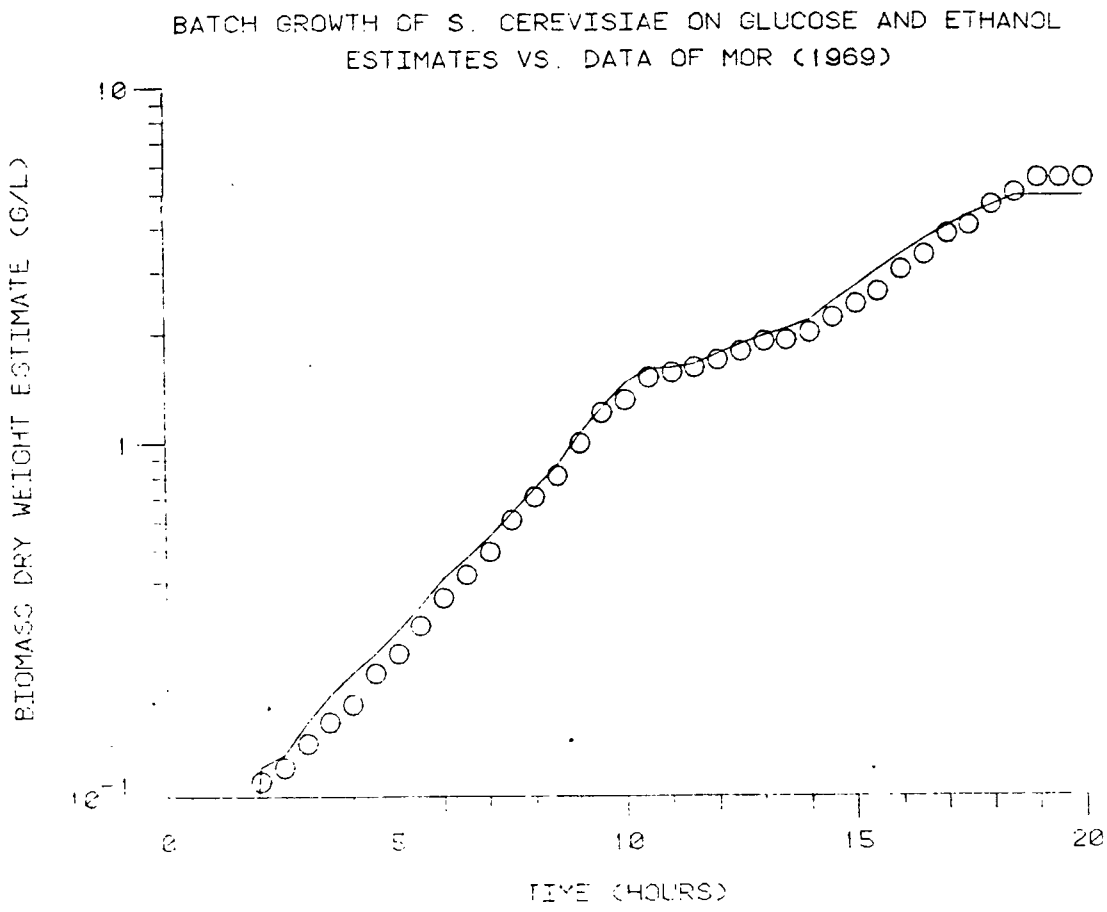
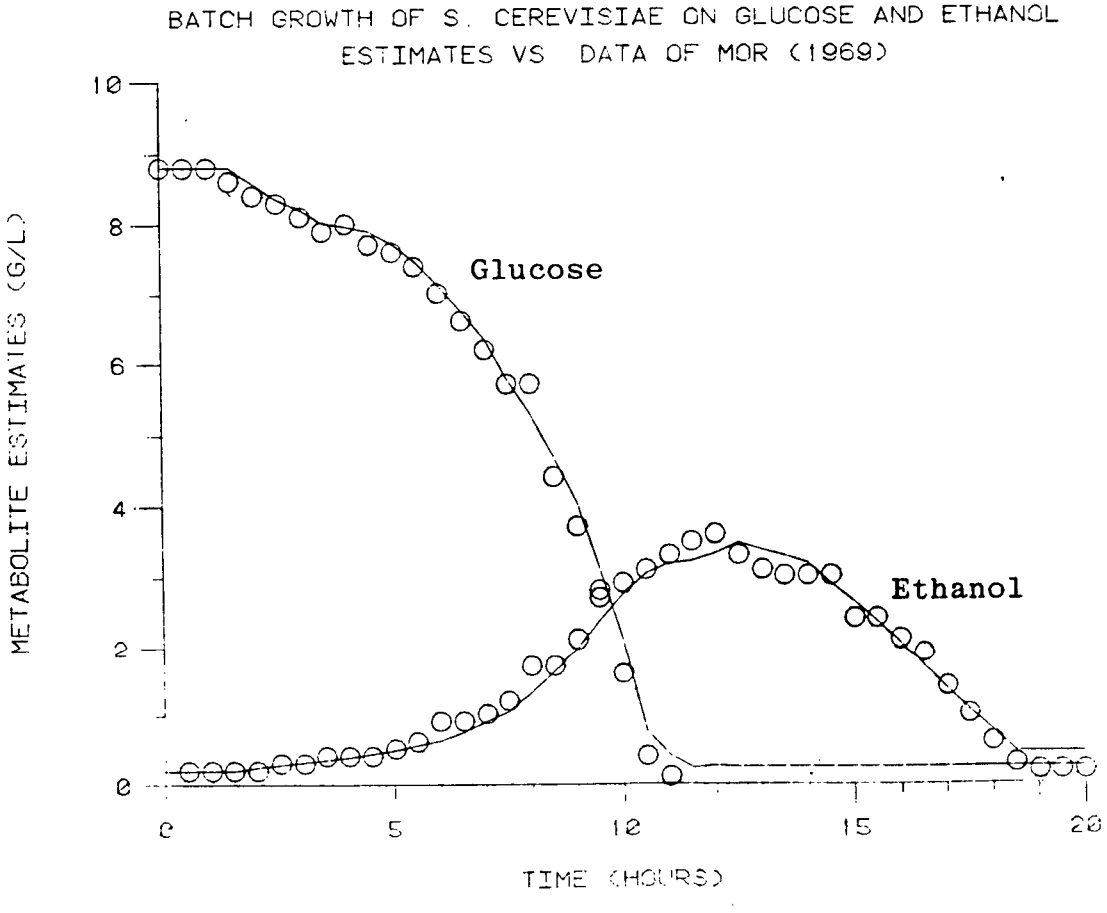


Figure 3 – Comparison of State Estimates with Analytical Data





Nomenclature

a	substrate stoichiometric coefficient
b	oxygen stoichiometric coefficient
c	ammonia stoichiometric coefficient
C	colored noise state
e	water stoichiometric coefficient
f	carbon dioxide stoichiometric coefficient
g	metabolite product stoichiometric coefficient
M_{CO_2}	molecular weight of carbon dioxide
M_{O_2}	molecular weight of oxygen
p	metabolite product concentration (g/l)
R_{CO_2}	carbon dioxide evolution rate (mmol/l/hr)
R_{NH_3}	ammonia addition rate (g/l/hr)
R_{O_2}	oxygen uptake rate (mmol/l/hr)
R_x	total growth rate of biomass (g/l/hr)
RQ	respiratory quotient
s	substrate (glucose) concentration (g/l)
s'	substrate (ethanol) concentration (g/l)
w_i	white noise
Y_{EtOH}	yield coefficient (g biomass/g ethanol)
Y_{gluc}	yield coefficient (g biomass/g glucose)
μ	specific growth rate of biomass
$'$	respiration phase variables
$\dot{y}$	first derivative of y with respect to time

Acknowledgment

The computer time for this project supported in part through the facilities of the Computer Science Center of the University of Maryland and in part through the facilities of the Chemical Process Systems Laboratory and the Systems Research Center of the University of Maryland.

References

1. Cooney, C. L., H. Y. Wang, and D. I. C. Wang. "Computer-Aided Material Balancing for Prediction of a Fermentation Parameter." *Biotechnol. Bioeng.*, **19**, 55-67, (1977).
2. Harmon, J., G. Lyberatos, and S. A. Svoronos. "Adaptive Optimization of Microbial Growth Processes." Paper WP10, ACC Meeting, Boston, MA, (1985).
3. Jazwinski, A. H. Stochastic Processes and Filtering Theory. New York: Academic Press, (1970).
4. Lundell, R. "Practical Implementation of Basic Computer Control Strategies for Enzyme Production." Computer Applications in Fermentation Technology, Third International Conference on Computer Applications in Fermentation Technology, Society of Chemical Industry, London, 181-189, (1982).
5. Mor, J. R. PhD Thesis, E.T.H. Zurich, Switzerland, (1969).
6. Mou, D. G. and C. L. Cooney. "Growth Monitoring and Control Through Computer-Aided On-Line Mass Balancing in a Fed-Batch Penicillin Fermentation." *Biotechnol. Bioeng.*, **25**, 225-255, (1983).
7. Rolf, M. J. and H. C. Lim. "Experimental Adaptive On-Line Optimization of Cellular Productivity of a Continuous Baker's Yeast Culture." *Biotechnol. Bioeng.*, **27**, 1236-1245, (1985).
8. Staniskis, J. and D. Levisauskas. "An Adaptive Control Algorithm for Fed-Batch Culture." *Biotechnol. Bioeng.*, **26**, 419-425, (1984).
9. Stephanopoulos, G. and K. Y. San. "Studies on On-line Bioreactor Identification - Theory." *Biotechnol. Bioeng.*, **26**, 1176-1188, (1984).
10. Svrcek, W. Y., R. F. Elliott, and J. E. Zalic. "The Extended Kalman Filter Applied to a Continuous Culture Model." *Biotechnol. Bioeng.*, **16**, 827-846, (1974).
11. Wang, H. Y., C. L. Cooney, and D. I. C. Wang. "Computer-Aided Bakers Yeast Fermentations." *Biotechnol. Bioeng.*, **19**, 69-86, (1977).
12. Ydstie, B. E. "Self-Tuning and Adaptive Control." Notes from Short Course in Advanced Process Control. University of Maryland, College Park, MD, (1985).
13. Ziegler, W. M. and A. E. Humphrey. ACS Meeting, Washington, D.C., (1979).