State Estimation and Error Diagnosis in Industrial Fed-Batch Yeast Fermentation

A. Hocalar
Pakmaya, P.K. 149, 41001 İzmit, Kocaeli, Türkiye; and Dept. of Electronic and Communication Engineering, Univ. of Kocaeli, İzmit, Kocaeli, Türkiye

M. Türker
Pakmaya, P.K. 149, 41001 İzmit, Kocaeli, Türkiye

S. Öztürk
Dept. of Electronic and Communication Engineering, Univ. of Kocaeli, İzmit, Kocaeli, Türkiye

DOI 10.1002/aic.10996
Published online September 13, 2006 in Wiley InterScience (www.interscience.wiley.com).

Measurements provide the basis for monitoring and control of industrial processes as well as model development and validation. Therefore, systematic approaches are of great value to increase accuracy and reliability of measurements. In bioprocesses, linear conservation relations such as elements and enthalpy can be employed to relate conversion rates. In this work, a systematic approach has been applied to production scale fed-batch yeast fermentations. The six data sets obtained from two industrial size bubble columns, one with 25 m^3 volumes and the other with 100 m^3 volumes, are analyzed for state estimation and error diagnosis. A statistical test is employed for error diagnosis. The serial elimination method is used to analyze and locate the source of errors. The conversion rates are calculated from primary measurements such as flow rates, temperatures, and concentrations. When available measurements are more than the degrees of freedom of the system, it is said that the system is redundant. The redundancy is, therefore, used for error detection and data reconciliation for the six data sets in this work. In addition to elemental balances, heat balance has been set up for the bubble columns, and metabolic heat production rate is employed as an additional measurement. The redundancy is employed for state estimation, and biomass concentration and specific growth rate have been estimated with great accuracy. The estimations can be further used for process monitoring and control. © 2006 American Institute of Chemical Engineers
AIChE J, 52: 3967–3980, 2006
Keywords: state estimation, data reconciliation, fed-batch, process monitoring, error diagnosis, baker's yeast

Introduction

In the fermentation industry, fermenters are operated for maximum efficiency and minimum production cost. In order to maintain competitiveness, the fermentations must be highly consistent, with minimum variation in product quality, maximum yield on raw materials, and minimum production of undesirable side products. The development of optimal and consistent product performance and attaining minimum batch to batch variation requires the careful control of both growth rate and carbon flow. In order to achieve a high degree of control, knowledge of the metabolic states and amount of biomass present in the fermenter are necessary.[1] In biotechnological processes, online and offline measurements form the

Correspondence concerning this article should be addressed to A. Hocalar at akifh@pakmaya.com.tr.

basis of information about the state of the process. The continued advancement in instrumentation technology has allowed careful measurement and control of a number of fermentation parameters such as pH, temperature, dissolved oxygen, and off-gas analysis.[1-3] However, these measurements have a limitation in providing information on the state of the culture.[4] The use of off-gas analyzers has led to direct determination of oxygen uptake rates, carbon dioxide evolution rates, and respiratory quotients. The metabolic state of baker's yeast fermentation, for instance, can be determined by monitoring the composition of exit gas in terms of oxygen and carbon dioxide.[1,5-10]

For monitoring, optimization, and control of cultivations, it is important to quantify the growth and metabolite formation, and this can only be done through measurement of the biomass and/or metabolite concentrations. The absence of reliable and cheap sensors for monitoring directly unmeasurable fermentation variables, such as biomass, specific growth rate, and substrate concentrations, are the main bottleneck for bioprocesses in an industrial environment.[1-3,11]

It is, therefore, necessary to transform primary data into derived data that are most identifying the process states.[11] Quantitative analysis of cultivation can be built up on these derived data. Although it is highly desirable to measure biomass concentration, it is either measured offline by withdrawing samples or inferred from other measurements. The classical method for biomass determination is the gravimetric method, in which the biomass concentration is quantified by its dry weight. Other biomass measurement techniques are in situ (online) optical density sensors (infrared light source, near infrared lamp, laser diode), capacitance based instruments, and ultrasonic and conductivity based instruments.[12,13] When primary measurements are transformed into conversion rates, easily derivable linear conservation equations, such as the chemical elements, charge, and enthalpy, can be used for this purpose. That is why estimation techniques developed in control theory have been applied for online estimation of bioprocess variables. This has induced the development of "software sensors," which associate a sensor (hardware) and estimation algorithm (software) in order to provide online estimates of unmeasurable variables and kinetic parameters. There are three classes of software sensors according to the models used by the sensor for estimation, and James et al.[3] have compiled lists of estimators, such as:

(a) mechanistic model based estimators
(b) black box model based estimators
(c) hybrid model based estimators.

Several mechanistic model based estimators have been published in the literature, such as for baker's yeast[8-10,14,15] and for penicillin.[16,17] Several articles have also been published on black box model estimators[18-20] and hybrid model based estimators.[21,22]

In combination with model-based calculations, they are used for monitoring as well as for manual and automatic process control. The validation of process models, as well as the determination of the values of model parameters, can be established only with the aid of measurements. Therefore, it is important to have an accurate and consistent set of measurements.[7] All measurements in the experiments more or less include errors, and techniques to detect and locate errors are valuable in the development of process modelling, monitoring, and control. The errors in the measurements arise from different sources and can be classified into four categories[23-25]:

- zero mean random noise, in agreement with accuracy specification
- sudden gross errors
- systematic measurement errors
- decreased measurement accuracy.

Random noises are unavoidable and can be removed by application of filtering such as the Kalman filter.[26] The reliability of data used in a model can be established by applying the usual methods of random error minimization, such as repeated experiments, multiple sensors, and careful calibrations. Data redundancy is another consideration for the validation of the actual measurements and mechanistic framework.[27] Redundancies can be employed for the systematic detection of the source of gross measurement errors and systematic measurement errors.

In this work, consistency analysis of primary measurements and estimation of biomass concentration from consistent measurements are aimed at by using the mechanistic modeling and redundant measurements obtained from production scale baker's yeast fermentation. The CO_2, O_2, and ethanol concentrations in gas phase, air flow rate, and substrate feed rate are measured, and then the consumption/production rates are derived from these primary measurements. Heat balance has been set up around the fermenters to obtain metabolic heat production rate and included as a redundant measurement, which is rarely available in the literature.[28-31] After the unmeasured conversion rates are calculated, the best estimate values for the measured conversion rates are obtained by recursive application of the reconciliation method.

Theory

Mechanistic process description

In the mechanistic process description, all cellular relations are lumped into a single average cell with constant composition (unstructured, non-segregated) and the cells exchanging material and energy with the environment. The description based on the simple mechanistic model gives the chance to check the consistency of experimental measurements. Mechanistic modeling enables us to establish linear relations between biochemical reactions by using conservation laws and reaction conversion rates. Although some of the reaction rates can be calculated from primary measurements, some cannot be measured directly or are not feasible to measure in an industrial environment.

General stoichiometry of a fermentation process can be written with respect to reaction rates as:

$$r_s CH_{s1}O_{s2}N_{s3} + r_n NH_3 + r_o O_2 \rightarrow r_x CH_{x1}O_{x2}N_{x3} + r_p CH_{p1}O_{p2}N_{p3} + r_c CO_2 + r_w H_2O \quad (1)$$

and the rate vector is

$$r = (r_s, r_o, r_n, r_x, r_p, r_c, r_w)^T \quad (2)$$

As seen in the stoichiometric equation, there are four constraints (C, H, O, N) and seven rates in the system. Degree of freedom is

defined as $F = M + N + 1 - I$, where M is the number of metabolic products, N is the number of substrates, I is the number of constraints, and 1 is for the biomass. According to the number of measurements, the systems can be classified as:

- underdetermined, number of measurements less than the degree of freedom;
- determined, number of measurements equal to the degree of freedom;
- over-determined, number of measurements higher than the degree of freedom.

Once the system is over-determined, the redundancy of the measurements can be used to calculate non-measured conversion rates, increase the accuracy of the available measurements through the application of least square calculation, identify the sources of gross measurement errors, identify the sources of inconsistencies in formulation of the mechanistic model, and improve the state estimation of the process.[23-25]

The elemental composition matrix for the fermentation process in Eq. 1 can be written as:

$$E = \begin{array}{c} \begin{matrix} \text{substrate} & \text{oxygen} & \text{ammonia} & \text{carbondioxide} & \text{ethanol} & \text{water} & \text{biomass} \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ s1 & 0 & 3 & 0 & p1 & 2 & x1 \\ s2 & 2 & 0 & 2 & p2 & 1 & x2 \\ s3 & 0 & 1 & 0 & p3 & 0 & x3 \end{bmatrix} \end{array} \begin{matrix} C \\ H \\ O \\ N \end{matrix} \tag{3}$$

According to the general conservation law, multiplying of the elemental composition matrices and conversion rate vector is zero:

$$E \cdot r = 0 \tag{4}$$

The conversion rate vector can be divided into measurable (r_m) and non-measurable (r_c) conversion rates. Similarly, the elemental composition matrix is also divided into measurable (E_m) and non-measurable (E_c):

$$E_m r_m + E_c r_c = 0 \tag{5}$$

If the number of measurable rates is equal to the degree of freedom (F), the E_c matrix is in quadratic form and easily inverted as:

$$r_c = -E_c^{-1} E_m r_m \tag{6}$$

In case the E_c matrix is not in quadratic form, the matrix inversion is calculated by Eq. 7 (pseudo inverse):

$$r_c = -(E_c^T E_c)^{-1} E_c^T E_m r_m \tag{7}$$

In the case where the system is over-determined, more realistic estimates of measured rates can be obtained from the balances in addition to the calculation of non-measured conversion rates. By inserting Eq. 7 in Eq. 5:

$$R r_m = 0 \tag{8}$$

$$R = E_m - E_c (E_c^T E_c)^{-1} E_c^T E_m \tag{9}$$

The matrix R is called the redundancy matrix,[23-25] and its rank specifies the number of independent equations in Eq. 7. If these dependent rows are removed from the redundancy matrix, we obtain rank(R) independent equations relating the K measured rates:

$$R_r r_m = 0 \tag{10}$$

where R_r is the reduced redundancy matrix, containing only the independent rows of R. If it is assumed that conversion rates contain errors originating from primary measurements, we can denote the true (r_m) and measured ($\bar{r}_m$) values, and error vector δ with:

$$\bar{r}_m = r_m + \delta \tag{11}$$

Here, the error vector is assumed normally distributed with mean value of zero and with variance-covariance matrix F:

$$E(\delta) = 0 \tag{12}$$

$$F \equiv E\left[(\bar{r}_m - r_m)(\bar{r}_m - r_m)^T \right] = E(\delta\delta^T) \tag{13}$$

In practice, all the measurements contain some errors and there is a residual in each part of Eq. 10:

$$\varepsilon = R_r \delta = R_r(\bar{r}_m - r_m) = R_r \bar{r}_m \tag{14}$$

$$E(\varepsilon) = R_r E(\delta) = 0 \tag{15}$$

Here, ε represents the residuals vector. Since the residuals have a mean of zero and the variance-covariance matrix for ε is given by Eq. 16:

$$P = R_r F R_r^T \tag{16}$$

The minimum variance estimate of the error vector (δ) is obtained by minimizing the sum of the squared errors scaled according to the level of confidence placed on the individual

Table 1. The Values of the Chi-Square Distribution[27]

Degrees of Freedom	Confidence Level					
	0.5	0.75	0.9	0.95	0.975	0.99
1	0.46	1.32	2.71	3.84	5.02	6.63
2	1.39	2.77	4.61	5.99	7.38	9.21
3	2.37	4.11	6.25	7.81	9.35	11.3
4	3.36	5.39	7.78	9.49	11.1	13.3
5	4.35	6.63	9.24	11.1	12.8	15.1

measurements. The solution of the minimization problem is given by:

$$\hat{\delta} = FR_r^T P^{-1}\varepsilon = FR_r^T P^{-1} R_r \bar{r}_m \tag{17}$$

where the hat ($\wedge$) in Eq. 17 specifies that it is an estimate. Using Eq. 17, we can find an estimate of the measured rates to be given by[32]:

$$\hat{r}_m = \bar{r}_m - \hat{\delta} \tag{18}$$

When the length of the residual vector differs significantly from zero, there must be a significant error in at least one of the measurements or in the system description. In order to quantify the residuals significantly different from zero, we use a statistical test introduced by Reilly and Carpani[33]:

$$h = \varepsilon^T P^{-1}\varepsilon \tag{19}$$

The test function h is given by the sum of the weighted squares of the residuals and the residuals are weighted according to their accuracy. When the raw measurements are uncorrelated, the test function h is chi-square (χ^2) distributed. The degree of freedom of the distribution is equal to rank(P) = rank(R). By comparing the h test values with χ^2 distribution, it is possible to detect systematic errors in the data set. The χ^2 distribution table at different confidence levels and different degrees of freedom is given in Table 1.

Serial elimination method

If h test values are higher than the given confidence level, one cannot conclude whether the unsatisfactorily large errors are due to systematic errors or large random errors.[24,34] One approach to eliminate measurement error at a given data set is serial elimination. The serial elimination method, first introduced by Ripps, attempts to find the source of an error by leaving out the balanceable rates one by one, followed by calculating the probability. By comparing the h test results, there is strong evidence for the presence of gross measurement errors in the measurement that is eliminated, if a significantly lower value is obtained.

Growth stoichiometry of yeast

To demonstrate the capabilities of the mechanistic process description, its application to industrial baker's yeast fermentation is described here. Baker's yeast is produced in aerobic fed-batch process. Metabolic behavior of the baker's yeast fermentation process can be defined in three different pathways using the following stoichiometric equations[29,30]:

Oxidative growth on carbon source:

$$CH_2ON_{0.0125} + Y^{ox}_{O/S}O_2 + Y^{ox}_{N/S}NH_3 \rightarrow Y^{ox}_{X/S}CH_{1.8}O_{0.5}N_{0.2} + Y^{ox}_{C/S}CO_2 + Y^{ox}_{W/S}H_2O \tag{20}$$

Fermentative growth on carbon source:

$$CH_2ON_{0.0125} + Y^{red}_{N/S}NH_3 \rightarrow Y^{red}_{X/S}CH_{1.8}O_{0.5}N_{0.2} + Y^{red}_{P/S}CH_3O_{0.5} + Y^{red}_{C/S}CO_2 + Y^{red}_{W/S}H_2O \tag{21}$$

Oxidative growth on ethanol:

$$CH_3O_{0.5} + Y^{eth}_{N/S}NH_3 + Y^{eth}_{O/S}O_2 \xrightarrow{\mu_e^{eth}} Y^{eth}_{X/S}CH_{1.8}O_{0.5}N_{0.2} + Y^{eth}_{C/S}CO_2 + Y^{eth}_{W/S}H_2O \tag{22}$$

Mixed metabolism can be shown as in Eq. 23:

$$CH_2ON_{0.0125} + Y_{O/S}O_2 + Y_{N/S}NH_3 \rightarrow Y_{X/S}CH_{1.8}O_{0.5}N_{0.2} + Y_{P/S}CH_3O_{0.5} + Y_{C/S}CO_2 + Y_{W/S}H_2O + Y_{Q/S} \tag{23}$$

The general metabolic equation can be written based on conversion rates:

$$r_sCH_2ON_{0.0125} + r_nNH_3 + r_oO_2 \rightarrow r_xCH_{1.8}O_{0.5}N_{0.2} + r_pCH_3O_{0.5} + r_cCO_2 + r_wH_2O + r_q \tag{24}$$

The conversion rate vector can be written as follows:

$$r = (-r_s, -r_n, -r_o, r_x, r_p, r_c, r_w, r_q)^T \tag{25}$$

In Eq. 24, the metabolic heat production is included in the stoichiometric equation. The degree of freedom of this process is three. Hence, at least three measured rates are needed in order to calculate the unmeasured conversion rates. Furthermore, in order to have best estimate values for the measured rates and to perform consistency analysis and serial elimination for conversion rates, we have to measure two more conversion rates than the degrees of freedom. For this purpose, heat balance is included in the elemental composition matrix. In this way, redundancy of the system is increased without changing the degree of freedom. Here, volumetric rates are derived from primary measurements, and the degree of freedom of this process is three (F = 3 + 4 + 1 − 5). The elemental composition matrix including metabolic heat can be written as:

$$E = \begin{array}{c} \begin{matrix} \text{biomass} & \text{ethanol} & CO_2 & H_2O & \text{glucose} & \text{oxygen} & \text{ammonia} & \text{heat} \end{matrix} \\ \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1.83 & 3 & 0 & 2 & 2 & 0 & 3 & 0 \\ 0.56 & 0.5 & 2 & 1 & 1 & 2 & 0 & 0 \\ 0.17 & 0 & 0 & 0 & 0.0125 & 0 & 1 & 0 \\ 560 & 683 & 0 & 0 & 467 & 0 & 383 & -1 \end{bmatrix} \end{array} \begin{array}{l} \text{carbon balance} \\ \text{hydrogen balance} \\ \text{oxygen balance} \\ \text{nitrogen balance} \\ \text{heat balance} \end{array} \quad (26)$$

By writing Eq. 4 for a new system description, one obtains:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1.83 & 3 & 0 & 2 & 2 & 0 & 3 & 0 \\ 0.56 & 0.5 & 2 & 1 & 1 & 2 & 0 & 0 \\ 0.17 & 0 & 0 & 0 & 0.0125 & 0 & 1 & 0 \\ 560 & 683 & 0 & 0 & 467 & 0 & 383 & -1 \end{bmatrix} [r_x \;\; r_e \;\; r_c \;\; r_w \;\; -r_s \;\; -r_o \;\; -r_n \;\; r_q]^T = 0.$$

The non-measured conversion rates could be calculated by separating the rate vector, and the composition matrix is separated into measurable and non-measurable by a simple matrix operation as shown below:

$$\begin{array}{c} \begin{matrix} \text{oxygen} & CO_2 & \text{ethanol} & \text{glucose} & \text{heat} \end{matrix} \\ \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 3 & 2 & 0 \\ 2 & 2 & 0.5 & 1 & 0 \\ 0 & 0 & 0 & 0.0125 & 0 \\ 0 & 0 & 683 & 467 & -1 \end{bmatrix} \end{array} \begin{bmatrix} r_o \\ r_c \\ r_e \\ r_s \\ r_q \end{bmatrix} + \begin{array}{c} \begin{matrix} \text{ammonia} & H_2O & \text{biomass} \end{matrix} \\ \begin{bmatrix} 0 & 0 & 1 \\ 3 & 2 & 1.83 \\ 0 & 1 & 0.56 \\ 1 & 0 & 0.19 \\ 383 & 0 & 520 \end{bmatrix} \end{array} \begin{bmatrix} r_n \\ r_w \\ r_x \end{bmatrix} = 0$$

$$\begin{bmatrix} r_n \\ r_w \\ r_x \end{bmatrix} = \begin{bmatrix} 0.8081 & 1.0829 & -0.2085 & 0.2163 & -0.0001 \\ -0.9033 & -1.0744 & -0.1408 & -0.4308 & 0.0016 \\ -0.5527 & -0.7406 & -1.0771 & -0.9818 & -0.0017 \end{bmatrix} \times \begin{bmatrix} r_o \\ r_c \\ r_e \\ r_s \\ r_q \end{bmatrix}$$

On the other hand, the redundancy matrix and reduced redundancy matrix for our system are obtained from Eq. 9 as follows:

$$R = \begin{bmatrix} -0.5527 & 0.2594 & -0.0771 & 0.0182 & -0.0017 \\ -0.3936 & -0.2554 & 0.1220 & -0.0097 & -0.0003 \\ 0.7872 & 0.5109 & -0.2440 & 0.0193 & 0.0006 \\ 0.7031 & 0.9421 & -0.4131 & 0.0422 & -0.0004 \\ 0.0012 & -0.0005 & 0.0001 & -0.0001 & 0.0000 \end{bmatrix}$$

$$R_r = \begin{bmatrix} -0.5527 & 0.2594 & -0.0771 & 0.0182 & -0.0017 \\ -0.3936 & -0.2554 & 0.1220 & -0.0097 & -0.0003 \end{bmatrix}$$

By means of the reduced redundancy matrix (Rr), the variance-covariance matrix for measurement error vector (F), and the variance-covariance matrix for residuals (P), the minimum variance estimate of the error vector (δ) is obtained. Finally, by subtracting these estimates from online measured rates, best estimate values are obtained for the measured conversion rates. These new values can be called reconciliated/balanced values.

Materials and Methods

All the measurements have been carried out in two technical scale bubble column fermenters, one with 25 m^3 volumes and the other with 100 m^3 volumes. The flows entering the fermenter are substrate (molasses), air, ammonia, acid, and antifoam agent. The carbon dioxide, oxygen (Servomex, 1400 B4 SPX), and ethanol (Vogelbuch GS 2/3) concentrations are measured in the gas phase in the exhaust line. Molasses and

Table 2. Primary Measurements and Conversion Rates

Primary Measurements	Units	Conversion Rates	Units	Symbol
Air flow	Nm^3/h	O_2 consumption rate	mol/m^3h	r_o
Substrate flow rate	m^3/h	CO_2 consumption rate	mol/m^3h	r_c
O_2 concentration	%	Ethanol prod./cons. rate	$Cmol/m^3h$	r_e
CO_2 concentration	%	Substrate consumption rate	$Cmol/m^3h$	r_s
Ethanol concentration	%	Metabolic heat prod. rate	kJ/m^3h	r_q
Liquid volume	m^3	Water production rate	mol/m^3h	r_w
Cooling water flow rate	m^3/h	Nitrogen consumption rate	mol/m^3h	r_n
Cooling water inlet temp.	°C	Biomass production rate	$Cmol/m^3h$	r_x
Cooling water outlet temp.	°C			
Fermenter temperature	°C			

Table 3. Sugar and Metabolite Concentrations for F1 Fermentation During Initial Period of Fermentation

t (min)	Sucrose (kg/m^3)	Glucose (kg/m^3)	Fructose (kg/m^3)
15	10.18	3.64	3.41
30	6.63	5.14	5.91
45	3.092	5.04	7.69
60	2.07	4.3	7.85
75	1.04	2.75	7.45
90	0.72	1.18	5.33
105	0.58	0.187	1.55
120	0.52	0.48	3.33

t (min)	Glycerol (kg/m^3)	Acetic Acid (kg/m^3)	Ethanol (kg/m^3)
15	0.31	0.19	0.70
30	0.05	0.13	1.54
45	0.35	0.22	3.33
60	0.45	0.29	4.07
75	0.64	0.31	6.46
90	0.72	0.18	7.59
105	0.83	0.21	9.75
120	0.82	0.25	9.59
240			8.93
360	0.04	0.09	8.08

ammonia flow rates are measured by electromagnetic flow meters (Krohne, IFM 090), and airflow is measured by a vortex type (EMCO, V-Bar 700) flow meter. The temperature and pH of the broth are controlled by a closed loop controller in PLC (Modicon, 612 03). The fermenter with 25 m^3 volumes is equipped with external half coils to remove metabolic heat; whereas in the 100 m^3 fermenter, an external plate heat exchanger is used to remove heat generated during fermentation. The schematic views of fermenters are given in Figure 1.

During fermentations, all influents are fed according to predetermined profiles. All the closed loop control operations are managed by PLC and process measurements collected via the SCADA system (Nematron, Paragon V5.4). All the data sets are collected in one-minute periods, but the reconciliation algorithm steps are run at 15 min periods. The primary measurements and derived and calculated conversion rates are given in Table 2.

Methodology

In order to calculate the unmeasured conversion rates, one first has to derive specific conversion rates from the primary measurements. Oxygen uptake rate (r_o) and carbon dioxide production rate (r_c) are determined as below using the inert gas balance[29,30]:

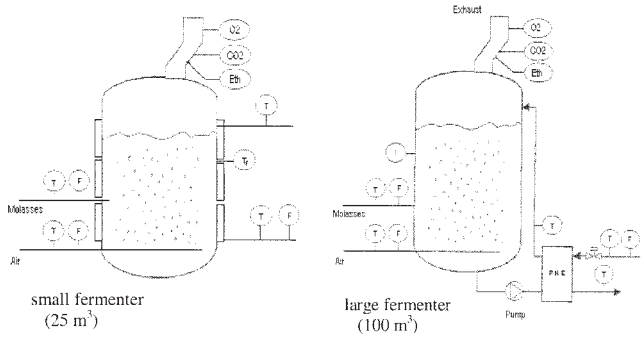

Figure 1. Views of fermenters and primary measurements.

$$r_o = \frac{F_n}{V}\left[\frac{P_{O_2}^{in}}{1 - P_{O_2}^{in} - P_{CO_2}^{in} - P_W^{in}} - \frac{P_{O_2}^{out}}{1 - P_{O_2}^{out} - P_{CO_2}^{out} - P_W^{out}}\right] \quad (27)$$

$$r_c = \frac{F_n}{V}\left[\frac{P_{CO_2}^{in}}{1 - P_{O_2}^{in} - P_{CO_2}^{in} - P_W^{in}} - \frac{P_{CO_2}^{out}}{1 - P_{O_2}^{out} - P_{CO_2}^{out} - P_W^{out}}\right] \quad (28)$$

The substrate consumption rate, r_s, is determined by measuring the molasses feed rate. However, in a fermentation where significant ethanol is formed (F1 fermentation), the substrate consumption rate is calculated from the dynamic mass balance by taking into account accumulation of sugars in the broth during the first 2 h of fermentation using the following dynamic equation:

$$r_s = \frac{F_i}{V_L} S_i - \frac{dS}{dt} \quad (29)$$

Accumulation of sugars and other metabolites during alcohol formation during the initial period of fermentation are shown in Table 3.

The ethanol production/consumption rate, r_e, is determined by online measuring of the alcohol concentration of the broth. Systematic errors that may occur due to neglect of the ethanol partition between liquid and gaseous phases are considered. In a fermentation where some ethanol has been formed by overdosing the carbon source, ethanol in the gas phase is taken into account in order to close elemental balances. Duboc and von Stockar[35] have studied ethanol partition between gas and liquid phases and derived the following equation at 30°C:

$$y_E = 0.532 x_E \quad (30)$$

Table 4. Quantification of Energy Sources in the Bioreactors

Feed:	$q_{feed} = \sum_{i=1}^{n} F_i \rho_i c_{pi}(T_i - T_f) = F_m c_{pm}(T_m - T_f) + F_{air}\rho_{air}c_{pair}(T_{air} - T_f)$
Metabolic:	$q_{metabolic} = r_q V_L$
Exchanger:	$q_{exchanger} = U_{exch}A_{exch}\Delta\bar{T}_L = F_w c_{pw}(T_{in} - T_{out})$
Surface:	$q_{surface} = h_{air}A_{sur}(T_{broth} - T_{surr})$
Evaporation:	$q_{evaporation} = F_{air}\Delta H_w^0(P_w^{out} - P_w^{in})$
Radiation:	$q_{radiation} = A_{sur}\varepsilon\sigma(T_f^4 - T_{surr}^4)$
Acid dilution and neutralization:	$q_{acid} = F_a\Delta H_d + F_a\Delta H_r$
CO_2 stripping:	$q_{CO_2} = F_c\Delta H_{CO_2}$

Table 5. Content, Mass, and Heat of Combustion of Each Component[30]

Element	Formula C-mol/mol	Mass (g/C-mol)	Heat Combustion kJ/C-mol
Biomass	$CH_{1.8}O_{0.5}N_{0.2}$	24.6	560
Glucose	$CH_2ON_{0.0125}$	30	467
Ammonia	NH_3	17	383
Ethanol	$CH_3O_{0.5}$	23	683

where x_E is the ethanol mole fraction in the liquid phase and y_E is the ethanol vapor content in the off gas. The gas phase ethanol mole fraction can be calculated from Eq. 30 when the liquid phase mole fraction is known. The metabolic heat production rate is measured using the dynamic energy balance around the fermenter as follows[29,30]:

$$\rho C_{p,b}\frac{d(VT)}{dt} = q_{feed} + q_{metabolic} + q_{exchanger} + q_{surface} + q_{evaporation} + q_{radiation} + q_{CO_2} + q_{acid} \quad (31)$$

The contribution of each term apart from $q_{metabolic}$ in Eq. 31 is calculated from online primary measurements using equations in Table 4, the details of which can be found in Türker.[29,30]

The metabolic heat production rate is calculated when the left-hand side of Eq. 31 is set to zero since the temperature of the broth is controlled at a set value. The compositions, mass weights, and heats of combustion of each component in the stoichiometric equation are given in Table 5.

The adjustments to the measured rates are related to the accuracy of the measurements and the structure of the equations. The relative errors for conversion rates are assumed to be 7% for r_o, r_c, and r_e, 4% for r_s, and 20% for the metabolic heat production rate r_q, based on our experience in our plant. The derivation of balanced/reconciliated conversion rates is performed in every programming cycle (15 min). The h test is applied only for the determination of confidence level at each programming cycle. Finally, more accurate biomass concentration values are obtained by using the reconciliated/balanced conversion rates.

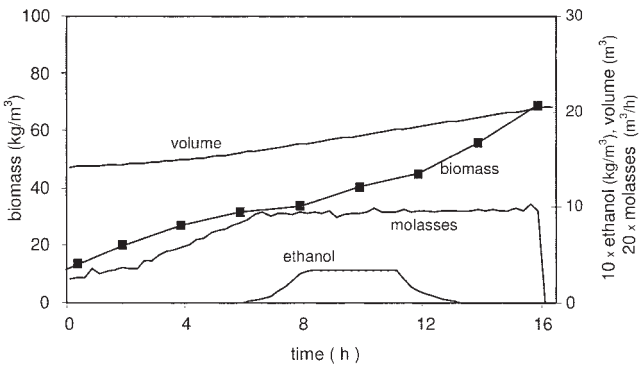

Figure 2. Progress of a typical fermentation obtained in the 25 m³ fermenter.

Straight lines for online measurements, —■— for offline biomass measurement.

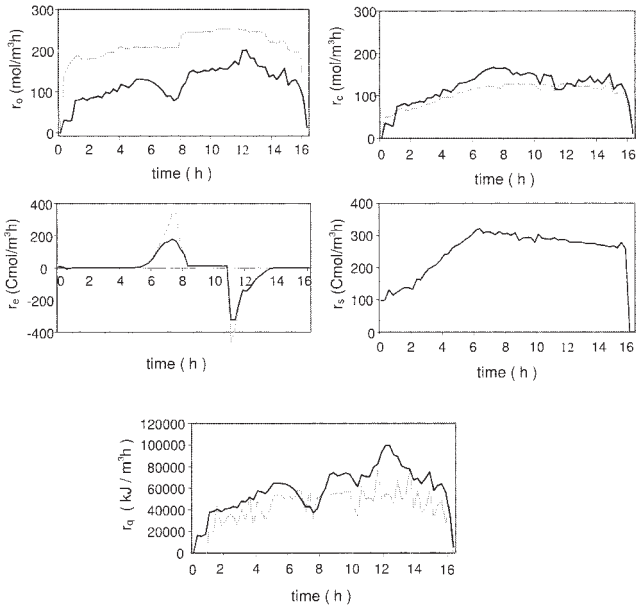

Figure 3. Online and reconciliated volumetric rates for b1 fermentation.

__ Online rates, ____ reconciliated rates.

Consistency Analysis Based on Carbon, Electron, and Energy Balances

We have analyzed the accuracy of online and reconciliated conversion rates, using elemental, electron, and heat balances, in addition to the h test. The balances can be written from Eq. 24 as follows[29,30]:

Elemental balances:

$$\begin{aligned} &C: && r_s = r_x + r_p + r_c \\ &H: && 2r_s + 3r_n = 1.8r_x + 3r_p + 2r_w \\ &O: && r_s + 2r_o = 0.5r_x + 0.5r_p + 2r_c + r_w \\ &N: && 0.0125r_s + r_n = 0.2r_x \end{aligned} \quad (32)$$

Electron balance:

$$\gamma_s r_s - 4r_o = \gamma_x r_x + \gamma_p r_p \quad (33)$$

Heat balance:

$$r_s\Delta H_S + r_n\Delta H_N = r_x\Delta H_x + r_p\Delta H_p + r_Q \quad (34)$$

Carbon (CR), electron (ELR), and energy (ER) recoveries can be defined as from Eqs. 32-34:

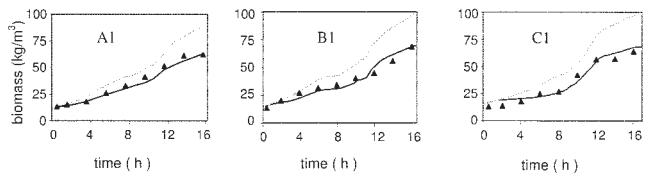

Figure 4. Online, reconciliated, and offline biomass estimation for typical pilot scale fermentations.

__ Online, ____ reconciliated, ▲ offline.

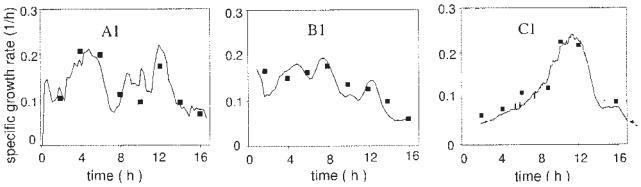

Figure 5. Reconciliated and offline specific growth rates estimations for typical pilot scale fermentations.

__ reconciliated, ■ offline.

$$CR: 1 = \frac{r_x}{r_s} + \frac{r_p}{r_s} + \frac{r_c}{r_s}$$

$$ELR: 1 = \frac{4r_o}{3.96r_s} + \frac{6r_p}{3.96r_s} + \frac{4.2r_c}{3.96r_s} \quad (35)$$

$$ER: 1 = \frac{Y_{X/S}(\Delta H_X - 0.2\Delta H_N + Y_{Q/X})}{\Delta H_S} + \frac{Y_{P/S}\Delta H_P}{\Delta H_S}$$

Results and Discussion

Derivation of reconciliated conversion rates in the small fermenter

The baker's yeast *Saccharomyces cerevisiae* was cultivated in fed-batch mode on commercial media based on molasses. The molasses was added initially at an exponential rate, then at a constant rate to the fermenter to maintain oxidative growth conditions. The progress of typical fermentation parameters is shown in Figure 2.

To evaluate how the model simulates the experimental data; the industrial size experiments are carried out in two fermenters: one with volume of 25 m^3 and the other 100 m^3 for industrial fermentations. Initially, the presented algorithm is applied to the data obtained from the smaller fermenter (25 m^3) for three different fermentations (A1, B1, and C1). The resulting online and reconciliated volumetric conversion rates, biomass estimations, estimated specific growth rates, estimated biomass yields, and carbon, electron, and energy recoveries are given in Figures 3-7, for a typical fermentation, respectively.

Figure 3 shows that online and reconciliated conversion rates closely match each other apart from the oxygen uptake rate, r_o. In order to analyze the source of errors in the measurements, we have applied the chi-square test to the B1 data set, and the results are shown in Table 6. The first column in Table 6 shows the magnitude of the chi-square test values obtained from online measurements. The magnitudes of the *h* test values in the first column of Table 6 are higher than the values in Table 1 for three degrees of freedom, indicating the presence of errors. We have applied the serial elimination method in order to locate the source of error in the measurements. The other columns of Table 6 are obtained by serial elimination of each conversion rate successively. The value of the test function is significantly reduced when oxygen uptake rate measurements are moved to the calculable part of the composition matrix; while other measurements are moved to the calculable part, the values of the *h* test remained high, indicating the presence of systematic errors in oxygen uptake rate measurements. The same fermentation has been performed in triplicate, and *h* test values are summarized in Table 7 for three data sets. For the A1 fermentation data set, *h* test values are low for the most part of the fermentation apart from a few points at the beginning and end of fermentation. For the C1 fermentation data set, *h* test values are reduced but still above the values from Table 1 when r_o is eliminated.

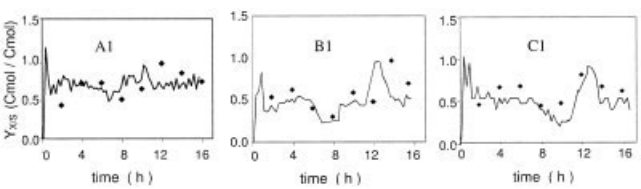

Figure 6. Biomass yields, $Y_{X/S}$ (Cmol/Cmol) from offline (♦) and reconciliated (__) data for A1, B1, and C1 fermentations, respectively.

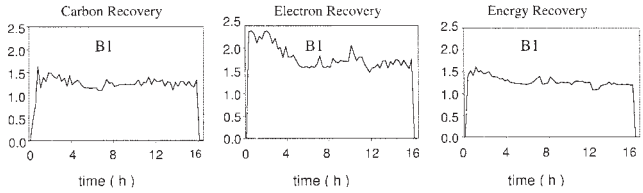

Figure 7. Carbon (CR), electron (ELR), and energy (ER) recoveries for online B1 fermentation data set, respectively.

In Figure 4, the offline biomass concentrations obtained along the fermentation are compared to online and reconciliated estimations for three data sets, and reconciliated biomass estimations give the best match to offline measurements. The estimated specific growth rate obtained from reconciliated data closely matches to that calculated from offline measurements, as shown in Figure 5, indicating that the specific growth rate can be estimated with acceptable accuracy from reconciliated process information. The reconciliated biomass estimation can also be used to get biomass yield, as shown in Figure 6.

Table 6. *h* Test Values when the Serial Elimination Method Is Applied for B1 Data Set

h					
	Serial Elimination of				
No Deletion	r_c	r_o	r_e	r_s	r_q
325.3	214.5	131.9	225.5	203.4	88.6
72.5	45.0	0.1	37.8	59.7	75.3
69.5	47.0	0.5	40.1	61.2	69.9
71.1	61.4	4.1	55.2	73.0	57.2
46.8	24.5	0.1	19.8	35.1	49.0
38.7	12.3	0.2	9.1	19.7	44.3
84.5	7.5	6.3	4.4	15.9	92.0
157.2	14.0	23.1	7.6	32.3	163.9
38.7	15.1	0.0	11.6	23.3	43.3
40.2	13.0	0.1	9.7	20.6	46.0
46.9	20.8	0.0	16.3	30.9	51.3
32.7	18.1	14.5	17.2	19.5	0.7
18.8	6.1	0.0	4.4	9.5	23.0
39.7	24.8	0.7	20.5	34.3	38.6
25.7	8.8	0.0	6.5	13.9	30.4
55.7	45.6	3.6	40.2	56.2	42.2

Table 7. *h* Test (Chi-Square) Values for the Experiments

			After Elimination of r_o		
A1	B1	C1	A1	B1	C1
50.2	325.3	168.5	20.3	131.9	59.1
39.0	72.5	50.8	12.3	0.1	13.3
50.6	69.5	53.0	17.3	0.5	11.6
27.2	71.1	27.2	5.7	4.1	2.3
17.9	46.8	42.8	0.9	0.1	9.7
27.4	38.7	38.0	2.8	0.2	11.0
27.7	84.5	35.5	2.6	6.3	8.7
27.2	157.2	40.6	4.6	23.1	10.3
21.6	38.7	31.0	5.7	0.0	7.8
25.7	40.2	33.1	7.2	0.1	2.5
21.3	46.9	65.4	5.0	0.0	32.3
22.9	32.7	47.0	5.9	14.5	27.8
23.9	18.8	12.4	6.2	0.0	2.5
16.1	39.7	24.7	1.7	0.7	10.9
33.0	25.7	40.5	10.5	0.0	11.6
15.9	55.7	42.6	10.9	3.6	15.6

The carbon, electron, and energy recoveries are presented in Figure 7, and significant deviation is observed in electron recovery. This observation is in line with that observed between online and reconciliated oxygen uptake rate measurements presented in Figure 3.

Derivation of reconciliated conversion rates in the large fermenter

The results obtained in the 100 m^3 fermenter are presented in Figures 8-10, respectively, for three sets of fermentations (D1, E1, and F1). In the D1 and E1 fermentations, little ethanol is formed at the beginning of fermentation; while in the F1 fermentation, some ethanol is generated by overfeeding the carbon source. The conversion rates calculated from online primary measurements are presented in Figures 8-10, together with reconciliated rates. Raw and reconciliated conversion rates seem to closely match each other for data sets D1 and E1. However, online metabolic heat production rates fluctuate and are not smooth, whereas reconciliated metabolic heat produc-

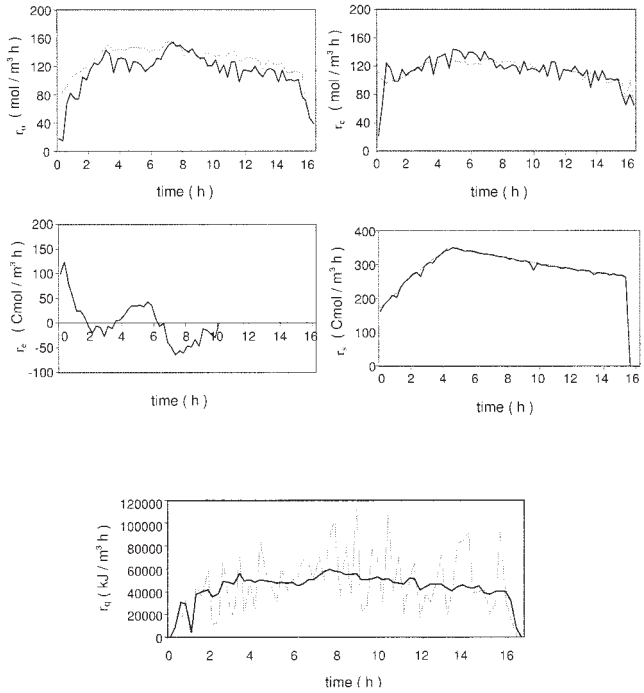

Figure 9. Online and reconciliated volumetric rates for E1 fermentation.

__ Online rates, ____ reconciliated rates.

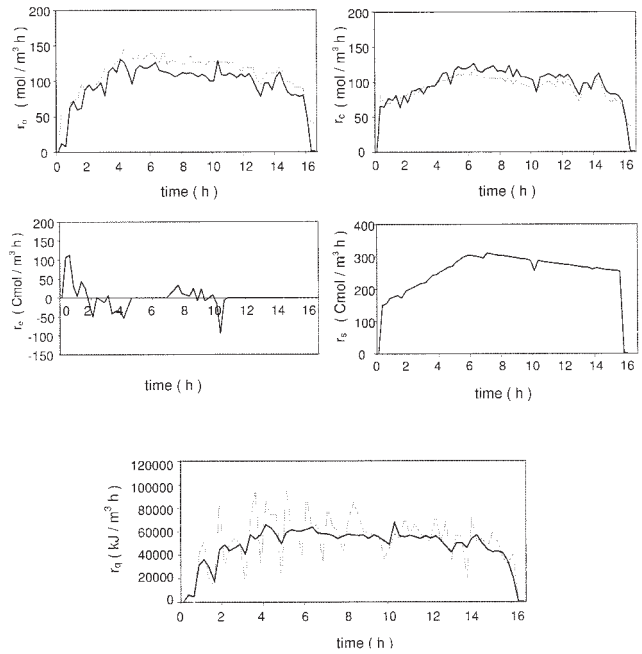

Figure 8. Online and reconciliated volumetric rates for D1 fermentation.

__ Online rates, ____ reconciliated rates.

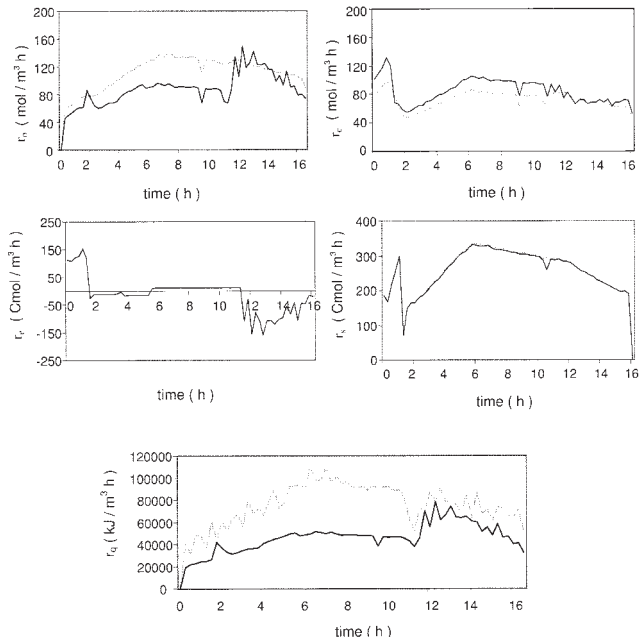

Figure 10. Online and reconciliated volumetric rates for F1 fermentation.

__ Online rates, ____ reconciliated rates.

Table 8. *h* Test (Chi-Square) Values for the D1, E1, and F1 Experiments

D1	E1	F1
1.6	0.0	0.0
10.0	9.3	68.4
3.2	12.6	31.9
1.0	20.3	33.5
5.7	3.8	32.1
6.3	7.8	28.7
4.2	4.4	25.7
3.5	0.2	28.7
8.0	0.5	29.3
0.8	6.4	28.3
3.4	12.2	28.2
4.6	34.5	13.0
7.6	13.4	3.0
2.3	3.2	1.7
10.7	27	3.1
4.3	23.3	3.8

tion rates are smooth as are other conversion rates. This is due to the dynamics of cooling water control valves.

However, in the F1 fermentation, online and reconciliated oxygen uptake, carbon dioxide production, and metabolic heat production rates do not match each other as closely as those in the D1 and E1 fermentations. All three online conversion rates diverge from reconciliated rates until the 12th hour, when ethanol consumption starts. Afterwards, online and reconciliated conversion rates converge.

In order to analyze the source of errors in the measurements, we have applied the chi-square test to all data sets (D1, E1, and F1 fermentation data sets), and the results are shown in Table 8. The magnitudes of the chi-square test values applied to the D1 and E1 data sets indicate that there is no suspicion about the quality of online measurements for these data sets apart from a few data sets that fail the test. Serial elimination has been applied to the E1 data set, and the results are shown in Table 9. Although the successive deletion of conversion rates from the matrix does not further reduce the magnitude of *h* test values compared to those in Table 1, the inconsistent data points yield the smallest performance indices only when the CO_2 production rate, r_c, is eliminated. However, the inconsistent data points still give higher performance indices when other conversion rates are eliminated. Therefore, we may conclude that online measurements of the E1 data set are generally reliable apart from the measurements at 3-4 and 11-13 hours of fermentation since *h* test values are higher than those in Table 1. Meanwhile, the magnitudes of *h* test values for F1 fermentation are higher than the values for three degrees of freedom until the 12th hour of fermentation; afterwards, *h* test values are significantly reduced, and online and reconciliated conversion rates converge to each other.

We have applied the serial elimination method in order to locate the source of error in the measurements in the F1 data set, and *h* test values are given in Table 10. When no deletion is made, the measurements fail the test until the last 4 hours of fermentation. The conversion rates are eliminated one by one in the calculation matrix. The *h* test values are significantly reduced when r_c, r_o, r_e, and r_s are eliminated, but not r_q. When r_q is eliminated, the *h* test values remained the same, indicating that r_q measurements are not faulty. We may conclude that systematic errors are not in a single conversion rate; rather, they may be distributed among several conversion rates. Therefore, the serial elimination method does not allow locating the

Table 9. *h* Test Values when the Serial Elimination Method Is Applied to E1 Data Set

h					
	After Deletion of				
No Deletion	r_c	r_o	r_e	r_s	r_q
9.3	0.0	4.9	3.5	7.6	13.5
12.6	3.3	9.8	8.4	12.3	3.6
20.3	5.9	14.9	13.3	17.9	3.7
3.8	0.1	1.0	0.6	1.6	7.0
7.8	0.0	3.1	2.1	5.0	11.8
4.4	2.1	0.1	0.3	0.1	9.3
0.2	0.3	0.0	0.0	0.0	1.1
0.5	0.6	0.5	0.6	0.9	0.0
6.4	1.7	3.7	3.1	4.4	0.7
12.2	3.0	8.6	7.4	10.7	2.9
34.5	10.8	26.3	24.3	30.0	5.1
3.2	0.0	1.1	0.6	1.6	5.9
8.6	1.1	6.3	5.1	8.4	5.2
2.1	0.3	0.3	0.1	0.4	5.2

Table 10. *h* Test Values when the Serial Elimination Method Is Applied to F1 Data Set

h					
	After Deletion of				
No Deletion	r_c	r_o	r_e	r_s	r_q
68.4	6.2	16.4	7.2	7.7	82.8
31.9	1.8	8.0	2.4	1.9	44.2
33.5	4.9	10.3	5.6	6.7	42.3
32.1	4.9	10.2	5.6	6.7	40.7
28.7	4.1	9.4	4.8	5.5	37.4
25.7	1.0	6.7	1.5	1.0	37.7
28.7	1.0	7.0	1.5	0.9	41.7
29.3	1.7	7.7	2.3	1.8	41.2
28.3	1.3	7.3	1.9	1.4	40.4
28.2	2.5	8.3	3.2	3.1	38.5
13.0	3.9	7.4	4.4	5.5	16.4
3.0	0.8	0.1	0.5	0.9	1.1
1.7	0.4	0.3	0.5	0.9	0.3
3.1	1.1	0.1	0.7	1.4	2.2
3.8	0.9	0.1	0.5	1.4	7.2

Table 11. *h* Test Values when the Serial Elimination Method Is Applied to F1 Data Set

h					
	After Deletion of				
No Deletion	r_c	r_o	r_e	r_s	r_x
244.8	242.4	217.7	8.2	101.2	81.2
54.2	0.0	19.0	52.7	41.5	45.4
70.4	3.6	43.1	67.3	38.1	43.6
59.1	1.7	30.0	57.8	37.0	42.0
68.3	7.1	45.7	64.4	32.9	38.7
45.1	0.0	14.3	42.6	35.4	38.9
48.4	0.0	13.2	44.9	39.3	43.0
42.4	5.2	1.5	28.1	41.1	42.4
41.7	5.2	1.2	26.5	40.4	41.7
40.1	2.5	3.8	31.0	37.7	39.7
23.9	0.8	11.7	23.0	14.9	17.4
8.0	7.4	6.2	1.7	1.8	1.5
4.3	3.0	4.1	3.9	0.1	0.1
7.8	2.0	5.5	6.6	1.7	2.8
8.1	2.1	0.0	3.3	7.3	7.9

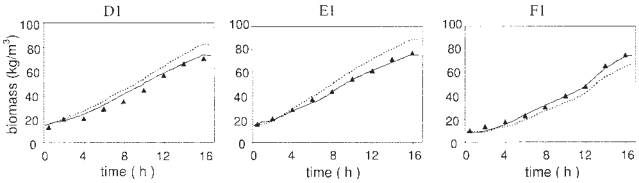

Figure 11. Biomass concentration obtained by mechanistic models and offline analysis.

Straight lines for reconciliated estimations, dashed line for online biomass estimation, ▲ for offline laboratory analysis.

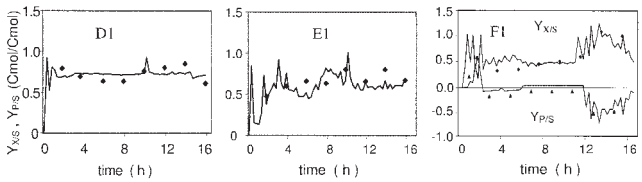

Figure 13. Biomass yields, $Y_{X/S}$ (♦), (Cmol/Cmol), and ethanol yields $Y_{P/S}$ (▲) from offline and reconciliated data (__) for D1, E1, and F1 fermentations, respectively.

source of errors if the errors are distributed among different measurements.

We have repeated the serial elimination for F1 fermentation by replacing r_q with experimentally measured r_x, and the results are shown in Table 11. In this case, *h* test values are significantly reduced only when the CO_2 production rate is eliminated during the whole period of the fermentation. When the oxygen uptake rate measurements, r_o, are eliminated, *h* test values are also reduced during the second half of the fermentation. They remained the same when other conversion rates were eliminated apart from the final 4 hours of fermentation, indicating that the CO_2 production rate during the whole period and the oxygen uptake rate during the second half of the fermentation may have a systematic error. This may also indicate an error in airflow rate measurements since oxygen uptake and carbon dioxide production rates are calculated as combinations of exit gas composition and airflow rate. However, we have observed that the data obtained from primary measurements are more accurate during the latter part of fermentations related to high conversion rates. It is not easy to detect the differences in the inlet and exit gas compositions, resulting in more likely faulty measurements. In addition, heat balance improves when higher conversion rates are obtained, because relative contributions of other heat sources become insignificant compared to metabolic heat.[29]

The offline biomass concentrations obtained along the fermentations (D1, E1, and F1) are compared to online and reconciliated estimations in Figure 11 for three data sets, and reconciliated biomass estimations give the best match to offline measurements. In all data sets, the reconciliated algorithm improves the estimation of biomass concentrations. Furthermore, the estimated biomass concentration could be used to calculate specific conversion rates such as the specific growth rate. The estimated specific growth rate obtained from the reconciliated biomass concentration is shown in Figure 12 and closely matches to that calculated from offline measurements,

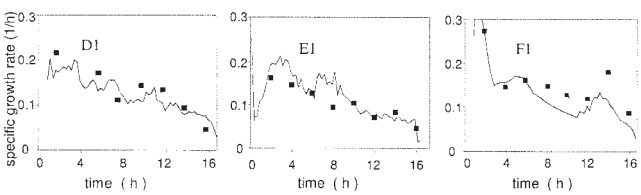

Figure 12. Reconciliated and offline specific growth rates estimations for typical industrial scale fermentations.

__ reconciliated, ■ offline.

indicating that the specific growth rate can be estimated with acceptable accuracy from reconciliated process information. Furthermore, the reconciliated process information can also be used to get biomass yields and ethanol yields on substrates, as shown in Figure 13.

In addition to the chi-square test, we have analyzed data in terms of carbon, electron, and energy recoveries. The carbon (CR), electron (ELR), and energy recoveries (ER) for D1, E1, and F1 fermentations are checked and plotted against time for online rates in Figure 14. Although initially the recoveries diverge at low metabolic activity, they are close to one, especially for D1 and E1 fermentations, as also justified by the chi square test. For the F1 fermentation data set, the carbon and electron recoveries are close to but less than one. However, energy recovery is close to one.

Replacing heat balance with degree of reduction balance

We have replaced heat balance with degree of reduction balance, since there is direct correlation between the degree of

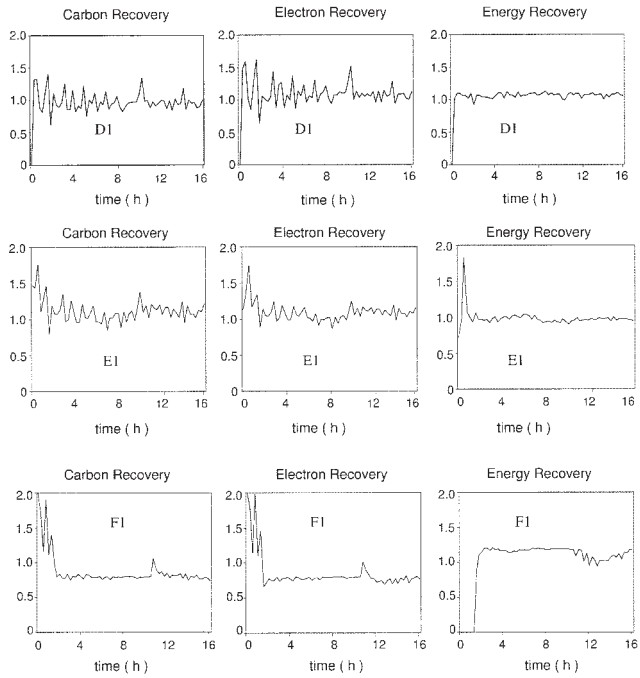

Figure 14. Carbon (CR), electron (ELR), and energy (ER) recoveries for online and reconciliated data for D1, E1, and F1 fermentations data sets, respectively.

Table 12. *h* Test Values when the Serial Elimination Method Is Applied to B1 Data Set Using Degree of Reduction Balance

h					
	After Deletion of				
No Deletion	r_c	r_o	r_e	r_s	r_x
198.7	3.6	165.3	107.5	64.9	67.9
120.0	6.6	92.0	82.3	55.7	57.6
126.2	2.5	103.4	81.4	51.2	53.5
117.0	0.0	103.5	72.7	41.2	43.8
58.7	0.0	64.1	59.3	35.4	37.5
34.6	6.4	12.3	37.2	33.5	33.9
83.5	209.8	128.8	1.7	83.5	70.4
141.8	340.7	289.0	9.7	154.4	125.5
38.3	2.7	27.5	42.6	32.2	33.2
43.3	2.6	33.2	46.4	34.1	35.2
53.5	2.3	44.1	54.0	37.9	39.3
167.1	86.3	200.8	49.7	0.0	0.6
88.7	15.8	127.6	50.5	14.9	17.6
60.0	0.8	70.8	51.4	27.4	29.5
39.8	0.0	40.6	35.3	22.0	23.2
46.7	0.0	51.2	49.9	30.5	32.3

reduction and the heat balance, and the degree of reduction balance is readily available from available measurements other than heat flux. The biomass production rate, r_x, can now be obtained from Eq. 33 since the rates are available online on the right-hand side of the equation as

$$r_x = \frac{\gamma_s r_s - 4r_o - \gamma_p r_p}{\gamma_x} \tag{36}$$

Now the system is over-determined and the degree of reduction balance can now be used for error diagnosis, data reconciliation, and state estimation. We have selected only three data sets out of six to test the feasibility and accuracy compared to those obtained in the previous section using heat balance. The chi-square test and serial elimination have been applied to the B1, D1, and F1 data sets only, and the results are shown in Tables 12-14, respectively. Normally, all three data sets contain errors,

Table 13. *h* Test Values when the Serial Elimination Method Is Applied to D1 Data Set when Degree of Reduction Balance Used Instead of Heat Balance

h					
	After Deletion of				
No Deletion	r_c	r_o	r_e	r_s	r_x
141.2	13.0	207.7	137.0	53.0	59.5
58.7	0.0	49.9	79.1	49.6	52.6
52.1	0.0	36.0	69.4	46.2	48.8
51.3	0.0	30.9	67.1	46.3	48.9
45.1	0.0	23.0	56.6	41.4	43.7
41.2	0.0	21.1	51.4	38.0	40.0
39.7	0.0	19.8	49.1	36.7	38.6
36.9	0.0	17.9	45.1	34.1	35.9
20.3	0.0	7.1	21.8	19.3	20.1
22.3	0.0	7.2	23.5	21.3	22.2
36.0	0.0	15.4	42.6	33.6	35.3
37.1	0.0	17.0	44.7	34.5	36.2
36.4	0.0	16.2	43.6	33.9	35.6
38.6	0.0	17.4	46.6	35.9	37.7
27.0	0.0	13.8	32.8	25.2	26.4
33.5	0.0	15.9	40.5	31.2	32.7

Table 14. *h* Test Values when the Serial Elimination Method Is Applied to F1 Data Set when Degree of Reduction Balance Used Instead of Heat Balance

h					
	After Deletion of				
No Deletion	r_c	r_o	r_e	r_s	r_x
61.5	363.7	385.5	99.1	54.5	32.4
81.8	87.1	65.9	3.6	2.5	1.4
30.7	9.6	56.9	50.0	19.9	22.8
28.5	0.1	18.8	36.9	35.7	27.3
24.8	7.6	41.2	42.2	17.8	20.4
27.1	1.7	4.7	25.2	26.8	27.1
28.6	3.8	2.7	24.0	28.9	28.8
31.5	3.7	3.3	27.0	31.7	31.7
31.2	3.7	3.9	27.7	31.3	31.4
30.6	3.7	3.5	26.7	30.7	30.7
29.1	3.8	3.2	25.2	29.3	29.3
9.8	41.3	61.6	19.1	0.0	0.1
209.6	186.9	297.3	77.6	0.5	0.1
203.0	179.3	290.2	75.3	1.7	0.1
195.0	147.7	272.9	72.9	0.4	2.1
220.5	114.2	259.3	74.9	4.9	8.5

having high *h* test values. When serial elimination is applied, it seems the carbon dioxide production rate, r_c, having measurement error in all data sets. The data reconciliation and state estimation have been applied, and the estimated and measured biomass concentrations and specific growth rates are presented in Figures 15 and 16, respectively. Biomass concentration and specific growth rate are not estimated with accuracy with online and reconciliated data for the B1 data set obtained in the small fermenter. However, the estimations in the large fermenter are better for both the D1 and F1 data sets, especially for the D1 data set. In fact, error diagnosis presented in the previous section also confirmed the accuracy of data set D1 using redundant metabolic heat measurements.

Conclusions

The combination of the estimation algorithm based on the mechanistic modeling and data reconciliation algorithm offers improved results for estimation of biomass concentration and, as a result, specific growth rate. By applying balancing techniques, more accurate conversion rates are obtained. Heat balance has been used as an additional measurement in the large scale. The advantages of metabolic heat measurement is that it is simple and could easily be applied with common process instruments normally available in industrial environments, such as flow meters and temperature sensors. The statistical test has failed, especially at the beginning of cultiva-

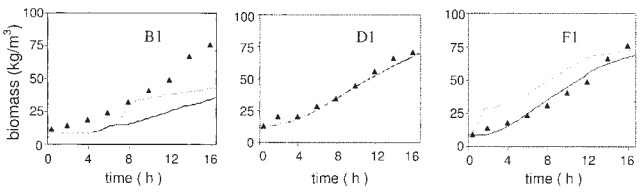

Figure 15. Online and reconciliated biomass estimations and comparison with offline biomass concentrations. __ reconciliated, ▲ offline.

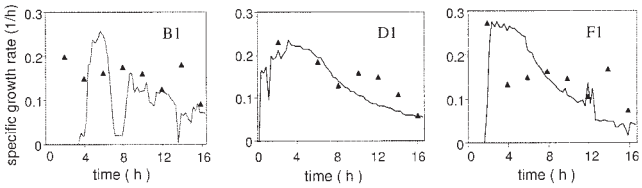

Figure 16. Offline and reconciliated specific growth rates.

tions where metabolic activity is low, whereas it improved when metabolic activity increased.

Generally, oxygen uptake and carbon dioxide production measurements seem to contain errors. This may be due to errors in the determination of exhaust gas composition, as well as in airflow measurements. The accuracy of exhaust gas composition is related to the difference between inlet and outlet gas compositions, which is also related to airflow rate. As the airflow rate increases, the difference between inlet and outlet gas compositions decreases, resulting in increased error in the determination of oxygen uptake and carbon dioxide production rates. This may be improved by reducing airflow rate and increasing agitation rate in mechanically agitated bioreactors. However, in bubble columns such as those used in this work, air not only provides oxygen but also supplies sufficient energy for mixing. In this work we had to use a relatively higher airflow rate in small bubble columns compared to the larger one to maintain oxidative conditions. This resulted in a lower difference between inlet and outlet gas compositions, as shown in Figure 17, resulting in reduced accuracy of the calculations. Duboc and von Stockar[35] have shown that the water content in the off-gas predominantly influences oxygen measurements. The calibration of airflow measurements is not an easy task, especially in industrial environments.

The redundant measurements with data reconciliation provide improved estimation of conversion rates. The biomass concentration and specific growth rate can be obtained from reconciliated data with great accuracy, as shown in this work. The direct determination of biomass concentration has been attempted, and sophisticated sensors have been developed in the literature. However, the use of biomass sensors is mainly limited to either clear liquors or dilute fermentation broths. As the trend moves towards using high cell density cultures, the quality and accuracy of online measurements improve, as presented here. For instance, the relative contribution of other heat sources compared to metabolic heat decrease; as a result, the accuracy of heat flux measurement increase as the cell density and volume of the bioreactor increase. Process engineers working in industry do not normally prefer to use sophisticated instrumentation for process monitoring and control. Therefore, common instruments combined with process knowledge can provide valuable information for error diagnosis, state estimation, and control of industrial bioreactors.

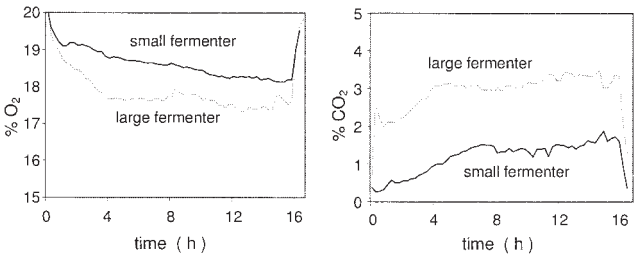

Figure 17. Comparison of exhaust gas compositions of small and large bubble columns.

Acknowledgments

The authors gratefully acknowledge the contributions of E. Tüller and O. Öztürk during the preparation of this work.

Notation

C_i = concentration of i (kg/m^3)
C_p = heat capacity (kJ/kg °C)
F_i = flow rate of i (m^3/h)
ΔH_x = heat of combustion of i (kJ/kg)
P_i = partial pressure of i (atm)
r_i = rate of consumption or production of i (C-mol/m^3 h)
r_q = volumetric heat production, (kJ/m^3 h)
$Y_{i/j}$ = yield of i over j
T = temperature (°C)
V = volume (m^3)
X = biomass (kg/m^3)
E_m = measurable part of elemental composition matrix
E_c = calculable part of elemental composition matrix
r_m = measurable part of conversion rate vector
r_c = calculable part of conversion rate vector
R = redundancy matrix
R_r = reduced redundancy matrix
F = variance-covariance matrix of error vector
P = variance-covariance matrix of residual vector
U = overall heat transfer coefficient (W/m^2 K)
A = area (m^2)
CR = carbon recovery
ELR = electron recovery
ER = energy recovery

Subscripts

a = acid
c = carbon dioxide
f = bioreactor
n = nitrogen
o = oxygen
s = substrate
p = product
x = biomass
T = transpose
w = water
$exch$ = exchanger
air = air
in = inlet
out = outlet
ox = oxidative
red = reductive
eth = ethanol
sur = surface
$surr$ = surrounding
$s1, s2, s3$ = hydrogen, oxygen, and nitrogen mole fraction in substrate
$p1, p2, p3$ = hydrogen, oxygen, and nitrogen mole fraction in product
$x1, x2, x3$ = hydrogen, oxygen, and nitrogen mole fraction in biomass

Greek letters

δ = measurement error
ε = residual vector
h = result of test function
γ = degree of reduction of I
μ = specific growth rate (h^{-1})
ρ = density of i, (kg/m^3)

Literature Cited

1. Cooney CL, Wang HY, Wang DIC. Computer aided material balancing for prediction of fermentation parameters. *Biotech Bioeng.* 1977; 19:55-67.
2. Dochain D. State and parameter estimation in chemical and biochemical processes: a tutorial. *J Process Control.* 2003;13:801-818.
3. James SC, Legge RL, Budman H. On-line estimation in bioreactors: a review. *Rev Chem Eng.* 2000;16(4):311-340.
4. Neeleman R. *Biomass Performance: Monitoring and Control in Bio-Pharmaceutical Production*. Ph.D. Thesis, Wageningen University, The Netherlands, 2002. ISBN 90–5808-733–6.
5. Aiba S, Nagai S, Nishizawa Y. Fed-batch culture of *Saccharomyces cerevisiae*: a perspective of computer control to enhance the productivity of baker's yeast cultivation. *Biotech Bioeng.* 1976;18:1001-1016.
6. Claes J. *Optimal Adaptive Control of the Fed-Batch Baker's Yeast Fermentation Process*. Ph.D. Thesis, Katholieke Universiteit Leuven, Belgium, 1999.
7. Keulers M. *Identification and Control of a Fed-batch Process*. Ph.D. Thesis, Technical University of Eindhoven, The Netherlands, 1993.
8. O'Connor GM, Sanchez-Riera F, Cooney CL. Design and evaluation of control strategies for high cell density fermentations. *Biotech Bioeng.* 1992;39:293-304.
9. Wang HY, Cooney CL, Wang DIC. Computer aided baker's yeast fermentations. *Biotech Bioeng.* 1977;19:69-86.
10. Wang HY, Cooney CL, Wang DIC. Computer control of baker's yeast production. *Biotech Bioeng.* 1979;21:975-995.
11. Hellinga C, Romein B, Luyben KChAM, Heijnen JJ. Error diagnosis and data reconciliation using linear conservation relations. In: van Impe JFM, Vanrolleghem PA (eds.). *Advanced Instrumentation, Data Interpretation and Control of Biotechnological Processes*. Amsterdam: Kluwer Academic Publishers; 1998:211-244.
12. Hatch RT, Veilleux BG. Monitoring of *Saccharomyces cerevisiae* in commercial baker's yeast fermentations. *Biotech Bioeng.* 1995;46: 371-374.
13. Olsson L, Nielsen J. On-line and in-situ monitoring of biomass in submerged cultivations. *Tibtech.* 1997;15:517-522.
14. Jobe AM, Herwig C, Surzyn M, Walker B, Marison I, von Stockar U. Generally applicable fed-batch culture concept based on the detection of metabolic state by on-line balancing. *Biotech Bioeng.* 2003;82(6): 627-639.
15. Zabriskie DW, Humphrey AE. Real time estimation of aerobic batch fermentation biomass concentration by component balancing. *AIChE J.* 1978;24:1, 138-146.
16. Mou DG, Cooney CL. Growth monitoring and control in complex medium: a case study employing fed-batch penicillin fermentation computer aided on-line mass balancing. *Biotech Bioeng.* 1983;25:257-269.
17. Mou DG, Cooney CL. Growth monitoring and control through computer aided on-line mass balancing in a fed-batch penicillin fermentation. *Biotech Bioeng.* 1998;325:225-255.
18. Karakuzu C. Türker M, Öztürk S. Modeling, on-line state estimation and fuzzy control of production scale fed-batch baker's yeast fermentation. *Control Eng Practice.* 2006;14:959-974.
19. Kitsuta Y, Kishimoto M. Fuzzy supervisory control of glutamic acid production. *Biotech Bioeng.* 1994;44:87-94.
20. Siimes T, Linko P, Numers C, Nakajima M, Endo I. Real time fuzzy knowledge based control of baker's yeast. *Biotech Bioeng.* 1995;45: 135-143.
21. Feyo de Azevedo S, Dham B, Oliveira F. Hybrid modelling of biochemical processes: a comparison with the conventional approach. *Computers Chem Eng.* 1997;21:751-756.
22. Schubert J, Simutus R, Dors M, Havlik I, Lübbert A. Hybrid modelling of yeast production processes. *Chem Eng Tech.* 1994;17:10-20.
23. Heijden Van der RTJM, Heijnen JJ, Hellinga C, Romein B, Luyben KChAM. Linear constraint relations in biochemical reaction systems: I. Classification of calculability and the balanceability of conversion rates. *Biotech Bioeng.* 1994;43:3-10.
24. Heijden Van der RTJM, Heijnen JJ, Hellinga C, Romein B, Luyben KChAM. Linear constraint relations in biochemical reaction systems: II. Diagnosis and estimation of gross errors. *Biotech Bioeng.* 1994;43: 11-20.
25. Heijden Van der RTJM, Heijnen JJ, Hellinga C, Romein B, Luyben KChAM. Linear constraint relations in biochemical reaction systems: III. Sequential application of data reconciliation for sensitive detection of systematic errors. *Biotech Bioeng.* 1994;44:781-791.
26. Stephanopoulos G, San KY. Studies on-line bioreactor identification I. Theory. *Biotech Bioeng.* 1984;26:1176-1218.
27. Stephanopoulos GN, Aristidou AN, Nielsen J. *Metabolic Engineering: Principles and Methodologies*. London: Academic Press; 1998.
28. Cannizzaro C, Valentinotti S, von Stockar U. Control of yeast fed-batch process through regulation of extracellular ethanol concentration. *Bioprocess Biosys Eng.* 2004;26:377-383.
29. Türker M. Measurement of metabolic heat in a production scale bioreactor by continuous and dynamic calorimetry. *Chem Eng Comm.* 2003;190:573-598.
30. Türker M. Development of biocalorimetry as a technique for process monitoring and control in technical scale fermentations. *Thermochimica Acta.* 2004;419:73-81.
31. Voisard D, Pugeud P, Kumar AR, Jenny K, Jayaraman K, Marison IW, von Stockar U. Development of a large scale biocalorimeter to monitor and control bioprocesses. *Biotech Bioeng.* 2002;80:125-138.
32. Nielsen J, Villadsen J. *Bioreaction Engineering Principles*. New York: Plenum Press; 1994.
33. Reilly PM, Carpani RE. Application of statistical theory of adjustment to material balances. 13th Congress Chem Eng Conf, Montreal, Quebec, Canada, 1963.
34. Wang NS, Stephanopoulos G. Application of macroscopic balances to the identification of gross measurement errors. *Biotech Bioeng.* 1983; 25:2177-2208.
35. Duboc P, Von Stockar U. Systematic errors in data evaluation due to ethanol stripping and water vaporization. *Biotech Bioeng.* 1998;58: 428-439.

Manuscript received Feb. 14, 2006, and revision received Aug. 8, 2006.