ISSN 0280–5316
ISRN LUTFD2/TFRT--5601--SE

Modeling and Control of a Continuous Stirred Tank Fermenter with Cell Recycle

Arvid Elmér

Department of Automatic Control
Lund Institute of Technology
June 1998

Department of Automatic Control
Lund Institute of Technology
Box 118
S-221 00 Lund Sweden

Document name
MASTER THESIS

Date of issue
June 1998

Document Number
ISRN LUTFD2/TFRT--5601--SE

Author(s)
Arvid Elmér

Supervisor
Sergio Valentinotti, Ulf Holmberg, EPFL, Lausanne,
Peter Niederberger, Nestlé, Lausanne,
Per Hagander

Sponsoring organisation

Title and subtitle
Modeling and Control of a Continuous Stirred Tank Fermenter with Cell Recycle

Abstract

A model for a CSTF with a cell recycle unit was identified in collaboration with Nestlé. Simulations of the model showed that it is possible to control the system in order to concentrate the biomass to 10-13 g/l with an overall dilution rate set to 1 1/h. The bacterium studied was *Lactobacillus johnsonii LJ1*, which produces lactic acid from lactose. The filter in the cell recycle unit was a cross-flow filter.

Key words
CSTF, cell recycle, cross-flow filter, modeling, identification, lactic acid control, yogurt

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title
0280–5316

ISBN

Language
English

Number of pages
34

Recipient's notes

Security classification

Modeling and Control of a Continuous Stirred Tank Fermenter with Cell Recycle

Master's Thesis done by

Arvid Elmér

Institut d'Automatique
Ecole Polytechnique Fédérale de Lausanne

June 1998

CONTENTS

1 Abstract

A model for a CSTF with a cell recycle unit was identified in collaboration with Nestec (Nestlé). Simulations of the model showed that it is possible to control the system in order to concentrate the biomass to 10-13 g/l with an overall dilution rate set to 1 h^{-1}. The bacterium studied was *Lactobacillus johnsonii LJ1* which produces lactic acid from lactose. The filter in the cell recycle unit was a cross-flow filter.

2 Introduction

Generally fermentation is referred to the production of alcohol or lactic acid from glucose by microorganisms (also called the cells or the biomass) in a fermenter. The studied bacterium LJ 1 produces lactic acid from lactose in an anaerobic process. There are many different types of fermenters in which the fermentation can take place. In a Continuous Stirred Tank Fermenter (CSTF), Figure 1, the fermentation can be maintained in exponential growth for a long time.

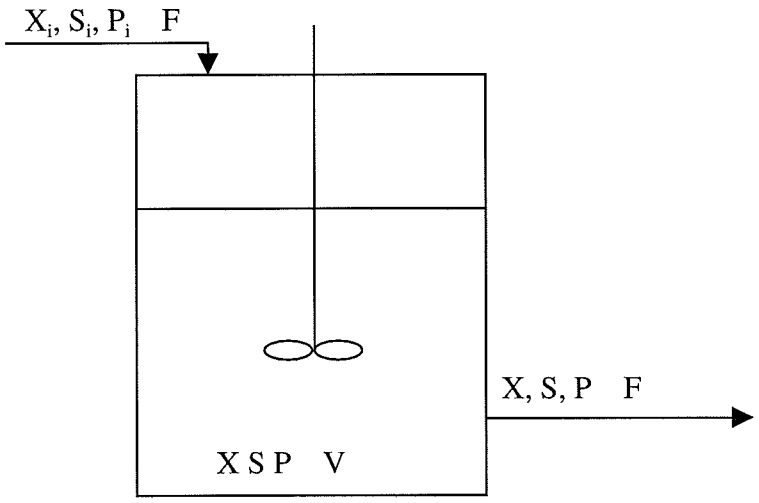


Figure 1. Schematic diagram of a Continuous Stirred-Tank Fermenter (CSTF). *X, S* and *P* are biomass, substrate and product concentrations respectively. The input flow concentrations are indexed *i*. *V* is the fermenter volume and *F* is the flow.

The mass balance for the biomass in a CSTF can be written as

$$FX_i - FX + Vr_X = V\frac{dX}{dt} \qquad (1)$$

where F is the flow, V the fermenter volume, X the biomass concentration [g/l] and r_X the *growth rate* [g/(l*hr)] of the cells. When the CSTF has reached steady-state conditions the change of biomass concentration with time equals zero ($dX/dt=0$) and eq. (1) becomes

$$\frac{F}{V} = \frac{r_X}{(X - X_i)} \qquad (2)$$

If the input flow is sterile ($X_i=0$) and the growth rate is written $r_X=\mu X$, eq. (2) becomes

$$\mu = \frac{F}{V} \equiv D \tag{3}$$

where D is called the *dilution rate* [h^{-1}] and μ is called the *specific growth rate.* Therefore, for a steady state CSTF with sterile feed, the specific growth rate μ is equal to the dilution rate D. For more details, look in any biochemical engineering book, for example Lee (1). The specific growth rate increases with an increase in the dilution rate until it reaches a maximum value corresponding to the maximum specific growth rate μ_{max}, which depends on the fermentation. If the dilution rate is increased beyond this value the cells will be washed out of the fermenter since the cell generation is less then the loss of cells with the outlet flow. The productivity is thus limited due to the loss of cells. One way of improving the productivity of the fermenter is to recycle the cells by separating them from the product stream using a filter. In this work a cross-flow filter was used where a rotor creates large shear rates in the fluid near the membrane which separate the cells. The shear field minimizes or prevents fouling of the filter. A CSTF with a cell recycle unit is shown in Figure 2.

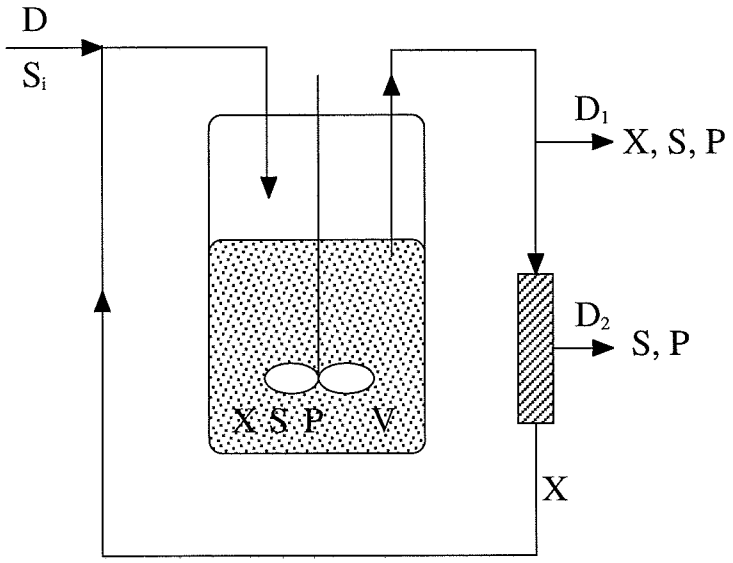


Figure 2. A CSTF with a cell recycle unit. D is the overall dilution rate, $D=D_1+D_2$. S_i is the substrate concentration of the input flow. The bleed dilution rate is noted D_1 and the dilution rate from the filter is called D_2. The biomass, substrate and product concentrations are represented as X, S and P respectively. V is the volume of the fermenter.

If all cells are recycled back into the fermenter the cell concentration will increase continuously and a steady state will never be reached. Therefore, to be able to operate the CSTF with cell recycle in a steady state mode, a bleed dilution rate D_1 from the fermenter is needed. The dilution rate D_2 represents the dilution rate from the filter. The overall dilution rate D is always equal to the sum of the two dilution rates leaving the system, $D=D_1+D_2$.

The *bleed ratio* γ is defined as

$$\gamma = \frac{D_1}{D} \tag{4}$$

When the bleed ratio $\gamma=1$, the cells are not recycled and $\mu=D=D_1$ in steady-state. However, when cell recycling is used, the bleed ratio γ is less then one and $\mu=\gamma D=D_1$. Since the specific growth rate μ is lowered with cell recycling, a higher overall dilution rate D can be applied and thus higher productivities without washing out the fermenter. Furthermore, high concentrations of lactic acid are growth inhibiting for the cells and therefore is cell recycling a way to concentrate the biomass and eliminate the toxic product. Normally it is the product from a fermentation what is aimed for in industry. But Nestlé want the increase the concentration of the biomass since they want to add the bacteria to their yogurts. The goal of this work was to show that it is possible to recycle the cells in a CSTF with a cross-flow filter in order to increase the production of biomass.

3 Methods

3.1 CSTF

The enriched medium that the CSTF was fed with is given in Table 1. The fermentation was performed under the anaerobic conditions given in Table 2. The inoculated cells were prepared according to standard procedures at Nestlé. The base composition added in order to have a constant pH is given in Table 3.

Medium components	**Concentration (g/l)**
Peptone de fibrine Nl	10
Yeast extract	15
Tween 80	1
Lactry	1
Lactose RAB	30

Table 1. Medium components

Reactor volume	3.5 l
Temperature	40°C
Stirring rate	100 rpm
pH	5.5
Inoculum	4%
Gas flow rate	CO_2 0.2 vvm N_2 0.2 vvm

Table 2. Fermentation conditions

Base components	
H_2O	3262.5 g
NaOH	391 g
KOH	846 g

Table 3. Base components

A sketch of the CSTF is shown in Figure 3. The input signals to the controlling computer system were the weight of the fermenter, pH, the dissolved oxygen (DO) and the gas composition. The weight was measured and regulated by the pump *P1* which pumped the output flow in order to have a constant volume. Th pH was measured and regulated the pump *P2* which added base to the fermenter to keep a constant pH. The base added was measured with a balance. The DO and gas compositions measurements were used to make sure that the fermentation was performed under the given gas conditions. The input flow was pumped by *P3* in order to achieve the given dilution rate *D*.

Samples were taken from a CSTF at five different dilution rates : 0.2, 0.35, 0.5, 0.7, 0.9 h^{-1}. A rule of thumb says that a CSTF has reached a steady-state after 4 residence times τ ($\tau=D^{-1}$). In order to confirm the steady-states, two samples for each dilution rate with an hour interval were taken after 5 residence times. The different dilution rates were applied in increasing order except for the dilution rate 0.2 h^{-1}, which was applied as the last dilution rate. To have some data to validate the model, samples were taken during the transient times of two step changes in the dilution rate: 0.2 – 0.7 and 0.2 – 0.5 h^{-1}.

The biomass concentration was determined by centrifuging 2x40 ml fermentation broth at 5000 rpm for 15 min and washing twice with TS solution followed by drying at 110°C on dried sea sand for 12 hr.

The lactose and lactic acid concentrations were determined with a HPLC (Hewlett Packard Serie 1100). After the injection of the sample (20 µl), the separation of the product was done on a Bio-Rad Aminex HPX-87H column (300x7.8 mm) using the following conditions: Temperature 35°C, Eluent H_2SO_4 (5 mM, flow rate 0.6ml/min), and a refractometer detector.

The titre of the culture was supposed to be determined by estimating the colony forming units (CFU) on MRS agar plates. Appropriate dilutions in sterile TS medium was prepared and poured on the plates. But the plates were contaminated so therefore no reliable conclusions could be made.

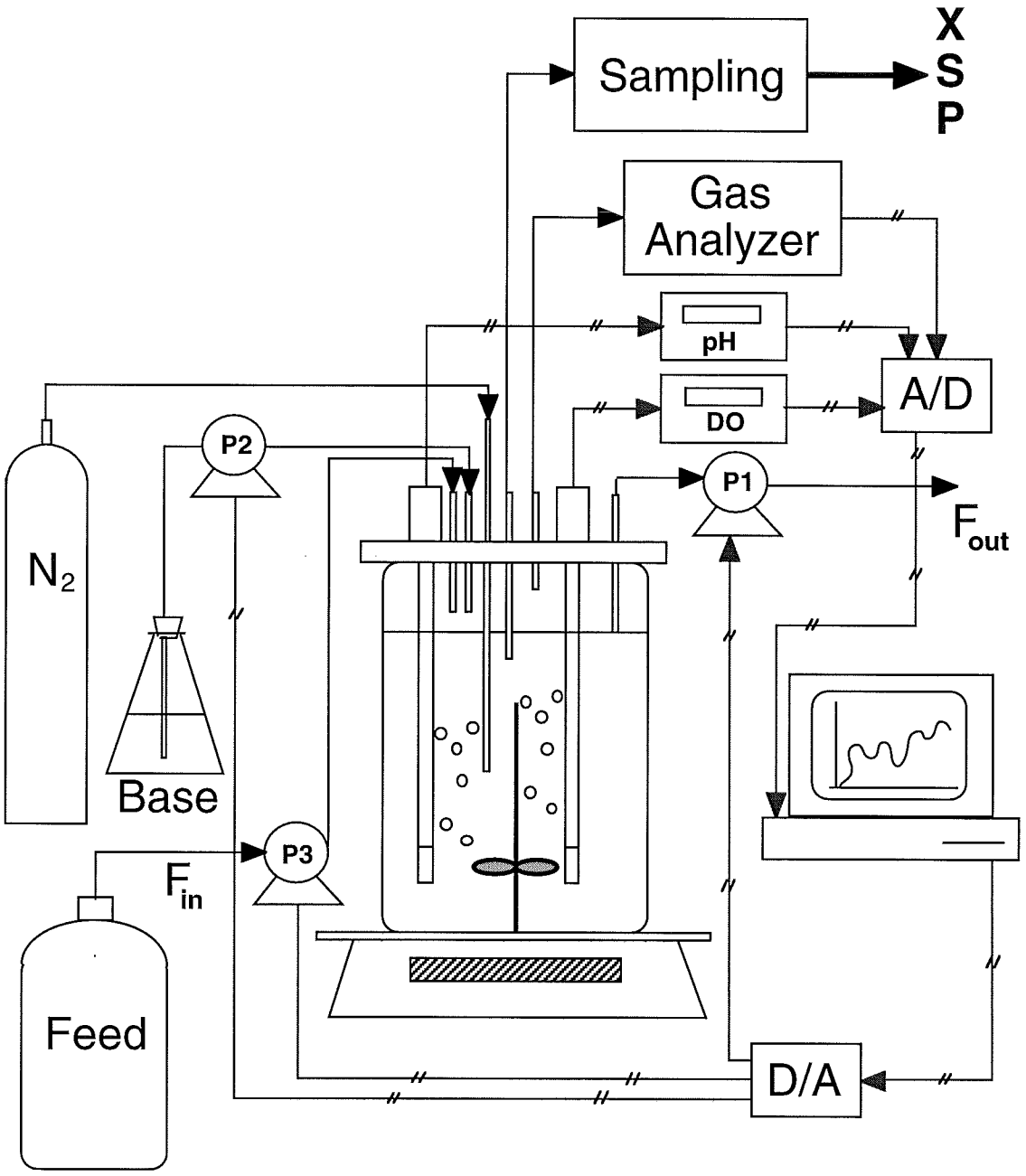


Figure 3. A sketch of the CSTF

3.2 The cell recycle unit

The cross-flow filter used to separate the biomass was a Mini-DMF unit from Pall Corporation. All the experiments were made with the rotor spinning at 2000 rpm. The fluid used as the fermentation liquid was water with a 10 g/l concentration of baker's yeast. No lactose or lactic acid was added since the particles are to small to effect the filter performance. A sketch of the cell recycle unit is shown in Figure 4.

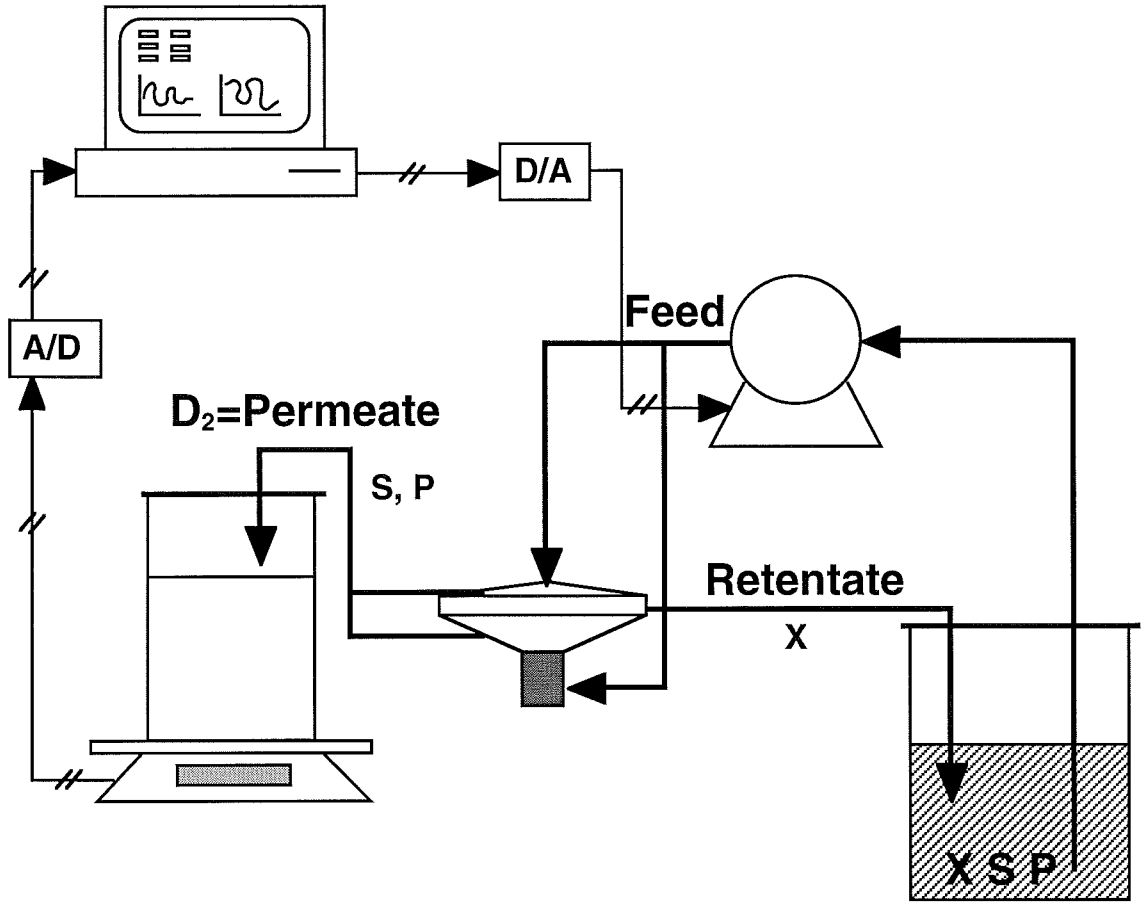


Figure 4. A sketch of the cell recycle unit

The fluid was pumped with an ALTIVAR 16 from the company SQUARE D. The pump was controlled from a PC with LabVIEW, a graphical programming software. In Figure 2 showing a CSTF with a cell recycle unit there are two streams leaving the system, D_1 and D_2 indicating that two pumps were needed for the cell recycle unit. However, only one pump was needed since the pump controlling the bleed D_1 was integrated in the control of the CSTF (Figure 3, pump *P1*). The dilution rate D_2 from the filter was measured by taking the time derivative of the weight. The balance used was a TP4KD from OHAUS.

The dynamics of a process can be modeled properly if the input signal excites sufficiently many modes of the process. A common way to obtain an input signal with many frequencies is to use a *Pseudo-Random Binary Signal* (PRBS). It is called pseudo-random because the signal is random within a defined period but in the next period the same alterations are made. For more details, look for example in Ljung (3). The input signal used for identification was a PRBS signal during two periods which altered the pump frequency between 10 and 13 Hz with a maximum impulse length of 28 seconds.

4 Modeling

Generally there are two different ways to obtain a mathematical model of a process, from prior knowledge, for example physical laws, or by doing experiments on the process. But usually it is not possible to make a complete model using physical laws. So therefore the approaches are often combined to determine some unknown parameters in the model. This is called *system identification*. The model of the CSTF was determined by formulating physical laws for the fermentation and then the unknown parameters were identified by experiments. The cell recycle unit was modeled only by experiments since no prior knowledge was given.

4.1 CSTF

The CSTF experiment was described with a model proposed by Taniguchi *et al.* (2). In their work they modeled a cultivation of lactic acid producing bacteria in a fermenter with cross-flow filtration. The model considers that the specific growth rate μ is a function of the product concentration P.

$$\mu(P) = \mu_{max} \exp(-\lambda P + \phi) \tag{5}$$

In the proposed expression the parameters λ and ϕ depend on P but in this work the model identification could be done with the parameters simplified to constants. The exponential term in (5) models the growth inhibiting effect of the product concentration. The following differential equations express the mass balance in a CSTF

$$\frac{dX}{dt} = (\mu(P) - D)X \tag{6}$$

$$\frac{dS}{dt} = D(S_{in} - S) - \overbrace{\left(\frac{\mu(P)}{Y_{X/S}(D)} + m\right)}^{rS} X \tag{7}$$

$$\frac{dP}{dt} = (\alpha\mu(P) + \beta)X - DP \tag{8}$$

X and S are the biomass and substrate concentrations. The substrate concentration is assumed not to be limiting the cell growth. The yield coefficients relating the substrate consumption and product formation to the biomass concentration are termed $Y_{X/S}$ and α. Normally $Y_{X/S}$ is treated as a constant, but experiments at Nestlé had shown a dependency on the dilution rate. The cell yield increased with the dilution rate. The parameters m and β represent that even if the cells are not growing, ($\mu(P)=0$), there are still substrate consumption and product formation. They are called the *maintenance coefficient* and *the non-growth associated production rate* respectively.

There are 7 parameters in this model: μ_{max}, λ and ϕ in eq. (5), $Y_{X/S}$ and m in eq. (7), α and β in eq. (8). The maximum specific growth rate μ_{max}, the yield α and the non-growth associated production rate β were provided by Nestlé from previous experiments. A third

degree polynomial for $Y_{X/S}$ was identified with least-square from the data provided. Nestlé had used other expressions to model the specific growth rate μ and the maintenance coefficient m. Therefore, the three parameters λ, ϕ and m had to be identified.

When doing an identification, a criteria is usually formulated to give a measurement of how well the proposed model fits the experimental data. The criteria of least squares is often used in identification problems. There are different versions but they are all based on the minimization of the sum of the squares of the errors. This sum is often called the *loss function* $J(\theta)$

$$J(\theta) = \sum_{k=1}^{N} \varepsilon_k^2 = \sum_{k=1}^{N} (y_k - \hat{y}_k)^2 \qquad (9)$$

where θ is the unknown parameter vector, N is the number of samples, ε_k is the error between the measured value y_k and the estimated value $\hat{y}_k$ at each sample k.

To be able to determine the unknown parameters, the identification was divided in two steps. First, the parameters λ and ϕ for the specific growth rate μ in eq (5) were identified by the least-squares method. Values for the product concentration P were taken from the samples. The theoretical values of μ at steady-states, i.e. μ equal the dilution rate D, eq. (3), were used as the measured value y_k.

With parameters λ and ϕ fixed, the next step was to identify the value of parameter m in eq. (7) by least-squares. Here the steady-state values of the term rS in eq. (7) were calculated from the samples and the value of μ was set to the corresponding dilution rate.

4.2 The cell recycle unit

The identification procedure was different for the cell recycle unit since no prior knowledge of the filter dynamics was given. There are many different identification methods such as frequency and transient-response analysis, correlation and spectral analysis and the general least-squares method to choose between. There is no method that is better than all the others. In this work the general least-square method was used for the cell recycle unit. A summarized description is given in *Appendix 1*

Let the system be described by

$$A(q^{-1})y(t) = B(q^{-1})u(t - n_k) + \varepsilon(t) \qquad (10)$$

This input-output model is called an ARX model were AR refers to the autoregressive part $A(q^{-1})y(t)$ and X to the extra input $B(q^{-1})u(t)$. The operators $A(q^{-1})$ and $B(q^{-1})$ are polynomials in the backward-shift operator q^{-1}, $q^{-1}y(t)=y(t-1)$. The delay between the input signal u and output signal y is called n_k. The term $\varepsilon(t)$ is a white noise disturbance, which in the ideal state contains all frequencies.

However, a disadvantage with the ARX model is that the disturbance term $\varepsilon(t)$ in reality may not be white noise. A C polynomial can be added to the model to compensate that.

$$A(q^{-1})y(t) = B(q^{-1})u(t - n_k) + C(q^{-1})\varepsilon(t) \qquad (11)$$

This model is called an ARMAX model where MA refers to the moving average $C(q^{-1})\varepsilon(\tau)$. A least-squares criterion can be obtained if the quantity to be minimized is taken as the sum of squares of the prediction error. An ARX model allows analytical solutions to the least-square criteria (9) but an ARMAX model demands the parameter estimation to be solved numerically.

An open loop identification can be done for processes which are stable. If the process is unstable the identification is usually done in closed loop. Since the cell recycle unit was stable, the identification was made in open loop for both the ARX and the ARMAX model. The determination of the model structure of the A,B and C polynomials are given in *Appendix 2.*

5 Controller design

It is difficult to get online measurements of the concentrations in a fermenter. Therefore, the control strategy was based on the product concentration P, since measurements of the product concentration were intended to be obtained online via calculations. A relation can be found between the base added to the fermenter in order to have a constant pH and the product concentration. From the given desired biomass concentration the corresponding product steady-state concentration was calculated via the model and set as the reference.

A block diagram of the control can be seen in Figure 5. Two RST controllers were used in cascade. The first RST controller determines the control signal D_2 i.e. the dilution rate from the filter. The second RST controller controls the pump in the Cell Recycle Unit (CRU) in order to get D_2 equal to the D_2*setpoint*

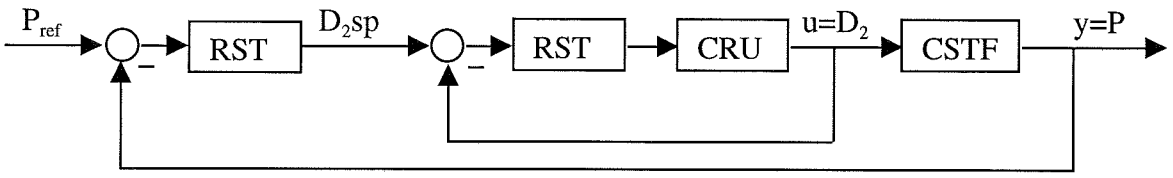


Figure 5. Control strategy. Two RST controllers in cascade

Both the fermenter and the cell recycle unit controller were designed with pole-placement. RST controllers were used instead of PID controllers since a RST controller has both feedback and feed forward control. A normal PID controller has only feedback control. A RST controller uses polynomials to place the closed loop poles in model given on an input-output form. Since the approach is taken from the literature (for example Åström (4)), only a summarized description is given here.

5.1 The fermenter controller

The model described by eq. (6)-(8) can be modified to include the filter with the filter efficiencies δ_X, δ_S and δ_P

$$\frac{dX}{dt} = \mu(P)X - \left(D_1 + (1-\delta_X)D_2\right)X \tag{12}$$

$$\frac{dS}{dt} = \left(D_1 + \delta_S D_2\right)(S_{in} - S) - \left(\frac{\mu(P)}{Y_{X/S}} + m\right)X \tag{13}$$

$$\frac{dP}{dt} = \left(\alpha\mu(P) + \beta\right)X - \left(D_1 + \delta_P D_2\right)P \tag{14}$$

The filter efficiencies are defined in the interval 0 to 1 where the value 1 implies ideal filtration and 0 no filtration. The filter efficiency for the cells is expressed as $(1-\delta_X)$ in eq (12) since ideal filtration means that no cells are lost with the dilution rate D_2 from the filter. The eq. (12) and (14) were linearized around the product concentration setpoint. Since the substrate concentration does not affect the product concentration, eq. (13) describing the substrate dynamics was neglected. The steady-state values can be derived as

$$\overline{X} = \frac{\left(\overline{D}_1 + \delta_P \overline{D}_2\right)\overline{P}}{\left(\alpha\mu(\overline{P}) + \beta\right)} \tag{15}$$

$$\mu(\overline{P}) = \left(\overline{D}_1 + (1-\delta_X)\overline{D}_2\right) \tag{16}$$

The variables with overbar represent steady-state values. Linearization of the model gives

$$\begin{bmatrix} \dot{\tilde{X}} \\ \dot{\tilde{P}} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \tilde{X} \\ \tilde{P} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tilde{D}_2 \tag{17}$$

$$y = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{X} \\ \tilde{P} \end{bmatrix} \tag{18}$$

where the time-invariant parameters are

$$
\begin{aligned}
a_{11} &= 0 \\
a_{12} &= -\lambda\mu(\overline{P}) \\
a_{21} &= \alpha\mu(\overline{P}) + \beta \\
a_{22} &= -(\lambda\alpha\mu(\overline{P})\overline{X} + \overline{D}_1 + \delta_P\overline{D}_2) \\
b_1 &= \delta_X\overline{X} \\
b_2 &= (1-\delta_P)\overline{P}
\end{aligned}
\tag{19}
$$

and $\tilde{X} = (X - \overline{X}), \tilde{P} = (P - \overline{P}), \tilde{D}_2 = (D_2 - \overline{D}_2)$ are the discrepancies from the steady-states. The above state-space system can be discretized and a discrete transfer function can be obtained

$$A_{ferm}(q^{-1})y(t) = B_{ferm}(q^{-1})u(t) + \varepsilon(t) \tag{20}$$

In this case the output y(t) is the measured steady-state discrepancy $\tilde{P}$ and u(t) is the input signal discrepancy $\tilde{D}_2$.

The fermenter should however face the whole system including the dynamics of the cell recycle unit. The achieved transfer function in (20) was therefore multiplied with the closed loop transfer function $H_{cl\_cru}(q^{-1})$ of the cell recycle unit to obtain a transfer function $H(q^{-1})$ of the whole system.

$$H(q^{-1}) = \frac{B(q^{-1})}{A(q^{-1})} = \frac{B_{ferm}(q^{-1})}{A_{ferm}(q^{-1})} * H_{cl\_cru}(q^{-1}) \tag{21}$$

The notation with the operator q^{-1} will be omitted for brevity since in the following all operators are also polynomials in q^{-1}. The control law

$$Ru = Tr - Sy \tag{22}$$

where R,S and T are polynomials, can be used to obtain a desired closed loop system as in Figure 6.

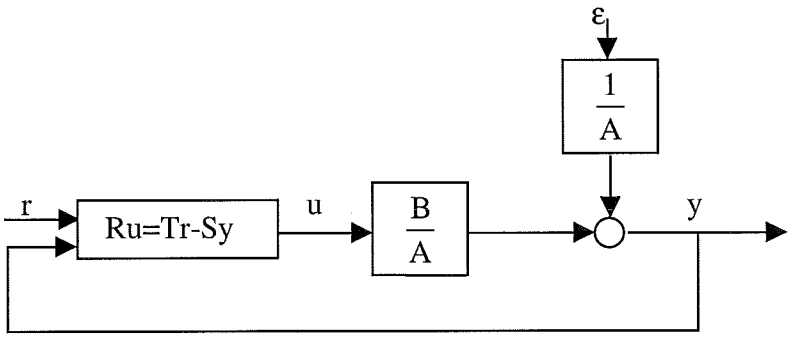


Figure 6. Block diagram of the closed loop system.

The reference signal r is zero in this case since as mentioned above, y is the steady-state discrepancy $\tilde{P}$. The desired closed loop characteristic polynomial A_m can be defined as

$$A_m = AR + BS \tag{23}$$

and thus can R and S be obtained from equation (23).
The closed loop polynomial A_m is chosen by specifying the roots (poles) which will in turn define the desired closed loop response. The polynomial T is normally calculated to get the right steady-state gain but since the reference signal r is zero in eq. (22), it has no use.

Integral action was chosen by including $(1-q^{-1})$ in the R polynomial. The coefficients in the R, S, and T polynomials depend on the desired steady-state the filter efficiencies and the dilution rate D since the linearization changes with them, and the desired closed loop polynomial A_m. An example of a pole-zero plot of the controller is shown in Figure 7 when the system was linearized around the product concentration 18 g/l, $D=1$, ideal filter efficiencies and the closed loop poles placed as in Figure 8. The coefficients of the controller polynomials are given in vector form in (24). The closed loop poles are also in Table 4. Two rules of thumb were used to place the closed loop poles:

- Not too far from the real axis to get the response damped and not too close to 1to avoid a to slow response
- Place the closed loop poles in order to have all the poles of the R polynomial on the positive half of the unit circle to avoid a ripple (oscillations) in the control signal.

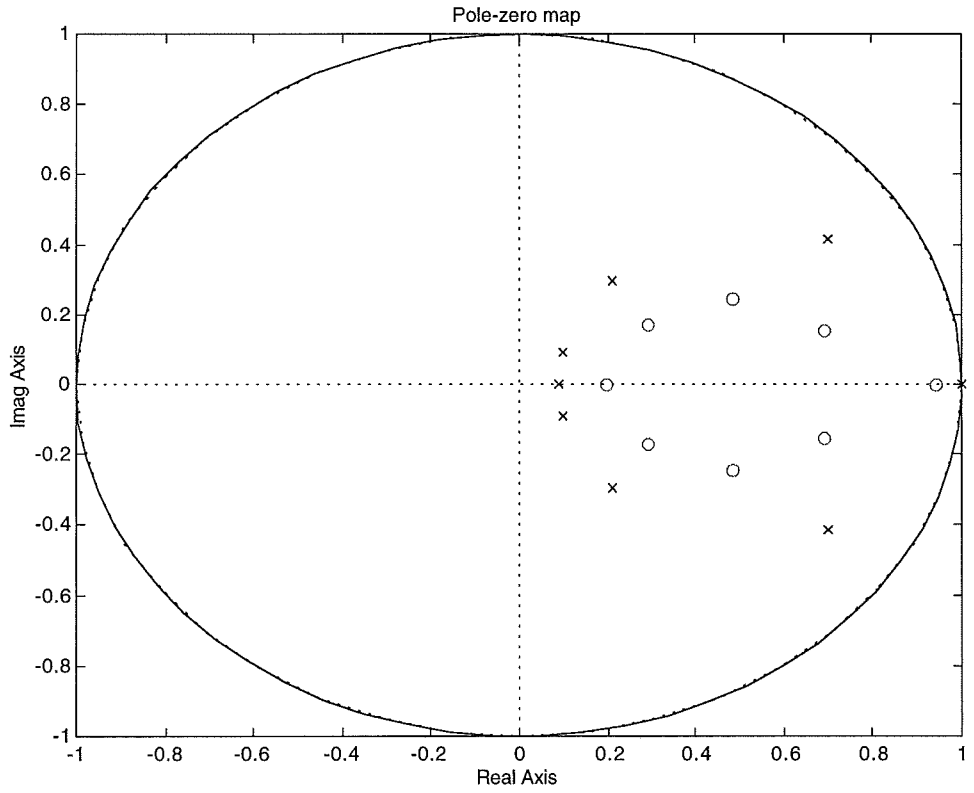


Figure 7 A pole-zero plot for the fermenter controller when $P_{sp}=18$ g/l, $D=1$, ideal filter efficiencies and the closed loop poles placed as in Figure 8. x=poles and o=zeros

$$
\begin{aligned}
R &= [1 \quad -3.1065 \quad 4.0429 \quad -2.8561 \quad 1.1883 \quad -0.3113 \quad 0.0464 \quad -0.0039 \quad 0.0001] \\
S &= [0.0785 \quad -0.3192 \quad 0.5595 \quad -0.5531 \quad 0.3378 \quad -0.1307 \quad 0.0313 \quad -0.0042 \quad 0.0002] \\
T &= [0.000078 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]
\end{aligned}
\tag{24}
$$

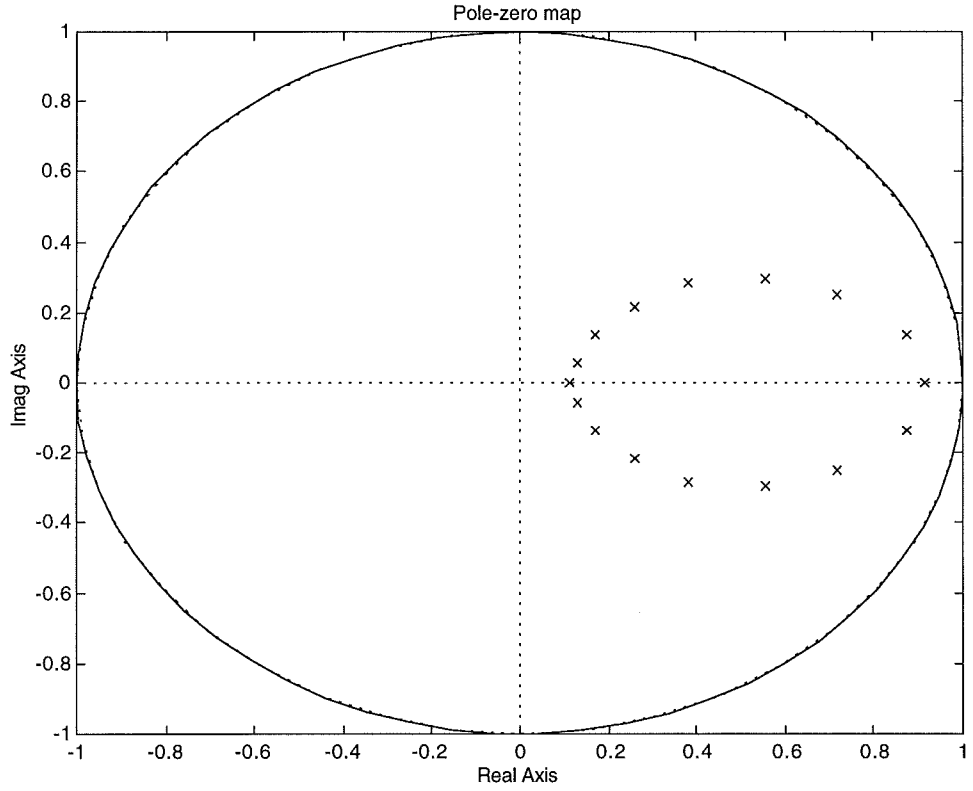


Figure 8 The closed loop poles for the whole system when the poles and zeros of the fermenter controller is placed as in Figure 7

Closed loop poles
0.9147
0.8761±0.1389i
0.7197±0.2513i
0.5541±0.2934i
0.3821±0.2823i
0.2591±0.2155i
0.1682±0.1342i
0.1298±0.0545i
0.1121

Table 4 The closed loop poles in Figure 8.

5.2 The cell recycle controller

The closed loop poles of the cell recycle unit were placed with the second RST controller. Integral action was chosen by including $(1\text{-}q^{-1})$ in the R polynomial. The polynomials $R,S,$ and T are given in (25) and the pole-zero plot of the controller is shown in Figure 9. The closed loop poles were placed with the same rule of thumbs as with the fermenter controller and they are shown in Figure 10. They are also given in Table 5.

$$\begin{aligned} R(q^{-1}) &= 1 - 1.1569q^{-1} + 0.3138q^{-2} - 0.1569q^{-3} \\ S(q^{-1}) &= 7.5740 - 12.3726q^{-1} + 9.0516q^{-2} - 2.5782q^{-3} \\ T(q^{-1}) &= 1.6747 \end{aligned} \tag{25}$$

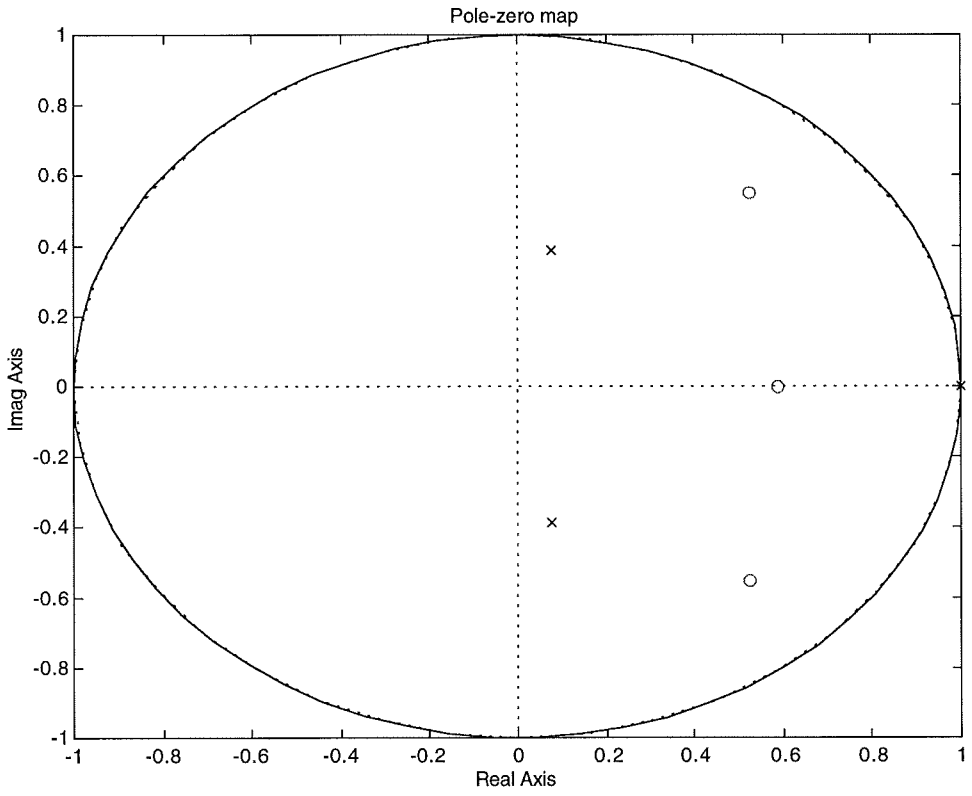


Figure 9. The pole-zero plot for the cell recycle controller. x=poles and o=zeros

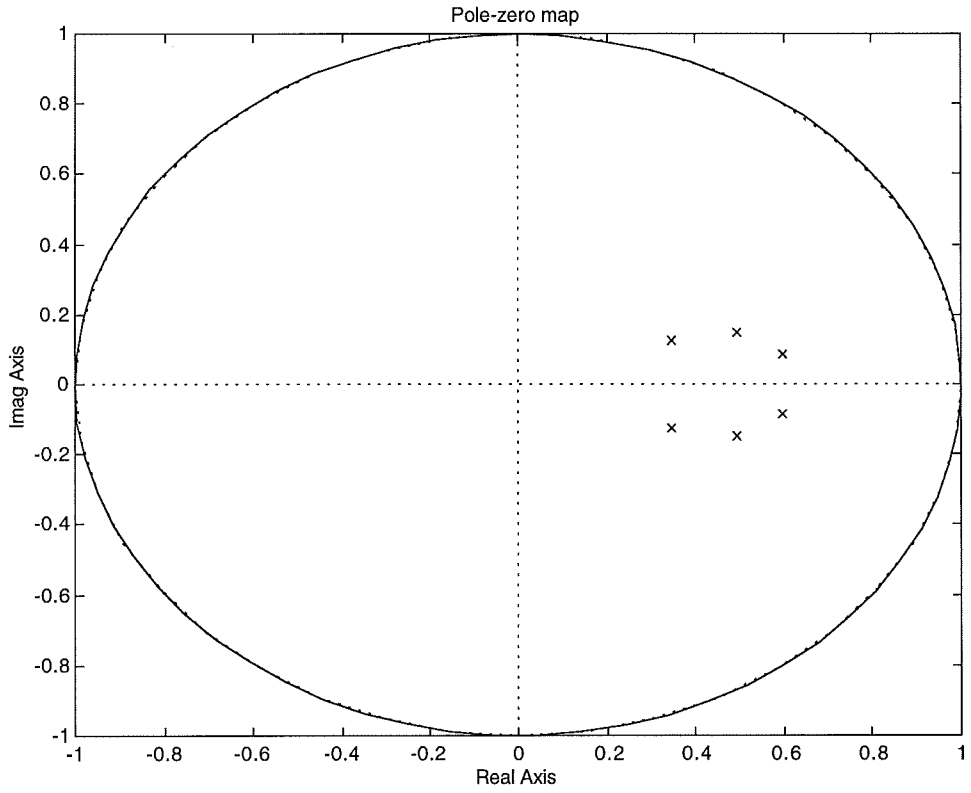


Figure 10. Closed loop poles for the cell recycle unit

Closed loop poles
0.5971±0.0864i
0.4950±0.1493i
0.3457±0.1257i

Table 5. The closed loop poles in Figure 10.

6 Robust control

6.1 The fermenter controller

The stability margins against an unstable response in closed loop can be calculated with the aid of frequency analysis. The Nyquist plot of the open loop transfer function for the whole system exemplified in chapter 5.1 is shown in Figure 11. The gain margin is 7.1 dB and the phase margin is 40.8° which allow some model uncertainty. A gain margin between 4 dB and 12 dB and a phase margin between 30° and 60° are generally accepted in practice. For more details, see Longchamp (5). The sensitivity function

$$S = \frac{1}{1 + H_0} \tag{26}$$

where H_0 is the open loop transfer function, models the transfer function between a perturbation on the output signal and the output. The sensitivity function curve of the whole system is given in Figure 12. The sensitivity function shows the desired behaviour with small values for high frequencies and big values for low frequencies since the integrator rejects the low-frequency disturbances.

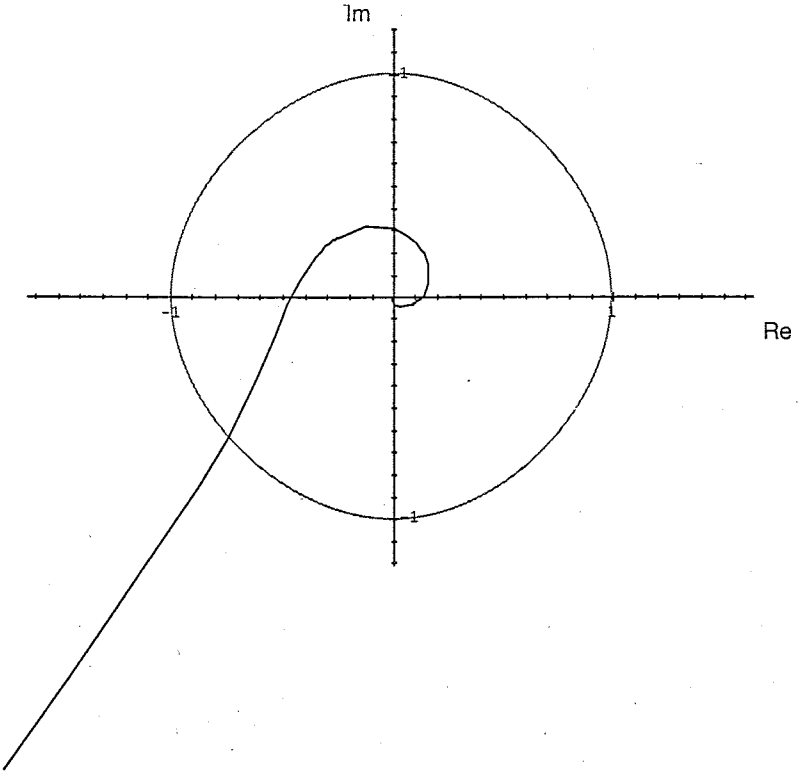


Figure 11. The Nyquist plot of the whole system exemplified in chapter 5.1

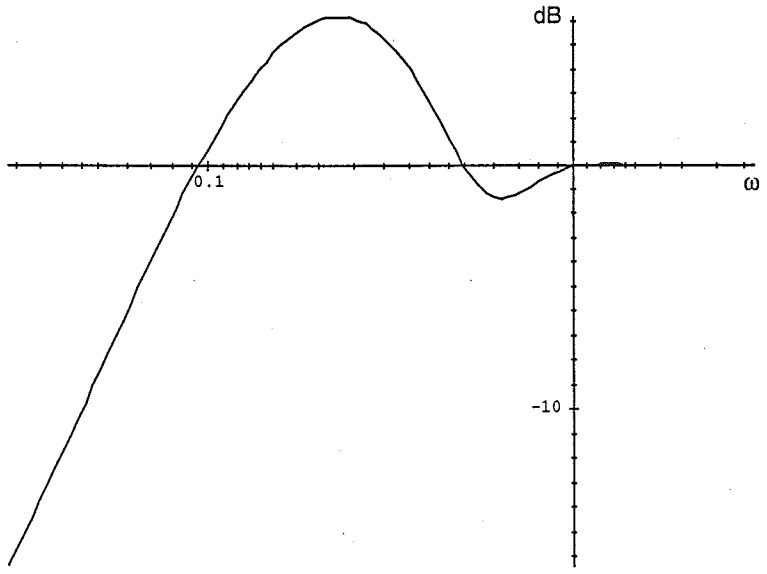


Figure 12. The sensibility function S in eq. (26) of the whole system exemplified in chapter 5.1

6.2 The cell recycle controller

The Nyquist plot of the open loop transfer function of the cell recycle unit is shown in Figure 13. The controller is given in (25). The gain margin is 11.6 dB and the phase margin is 65.9° which are big margins. The controller could thus have a faster response but this was not given a so high priority as for the fermenter controller since the cell recycle controller had a sampling time in seconds and the fermenter controller a sampling time in minutes. The sensitivity function is shown in Figure 14 which shows the same desired behaviour as the fermenter controller.

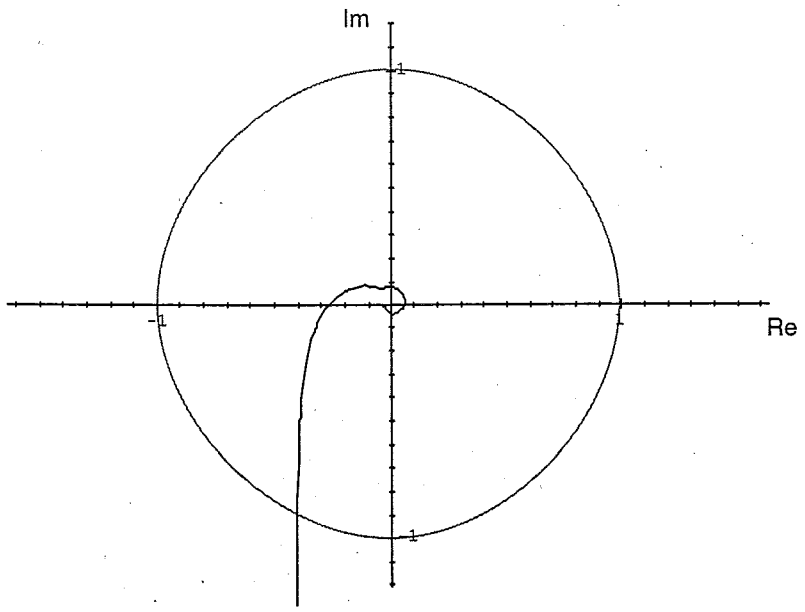


Figure 13. The Nyquist plot of the cell recycle unit

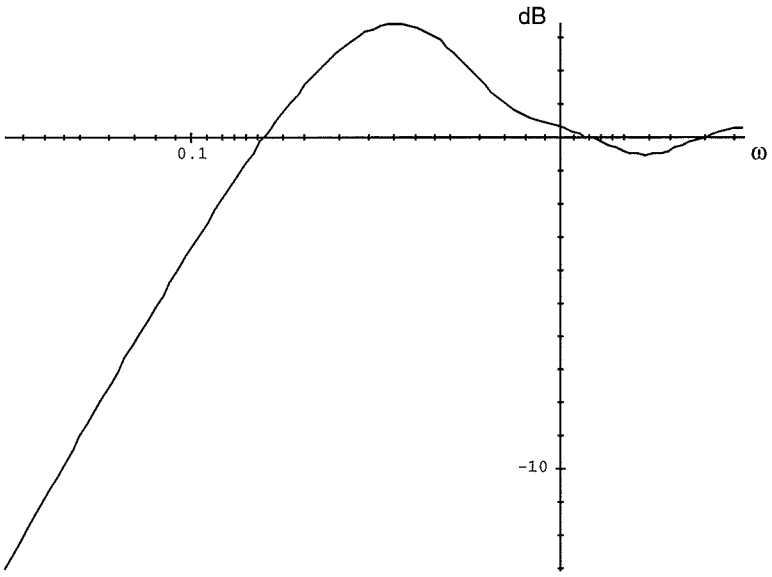


Figure 14. The sensibility function of the cell recycle unit

7 Results and Discussion

7.1 Model identification

7.1.1 CSTF

The least-squares fit for the yield $Y_{X/S}$ to the data provided by Nestlé is shown in Figure 15.

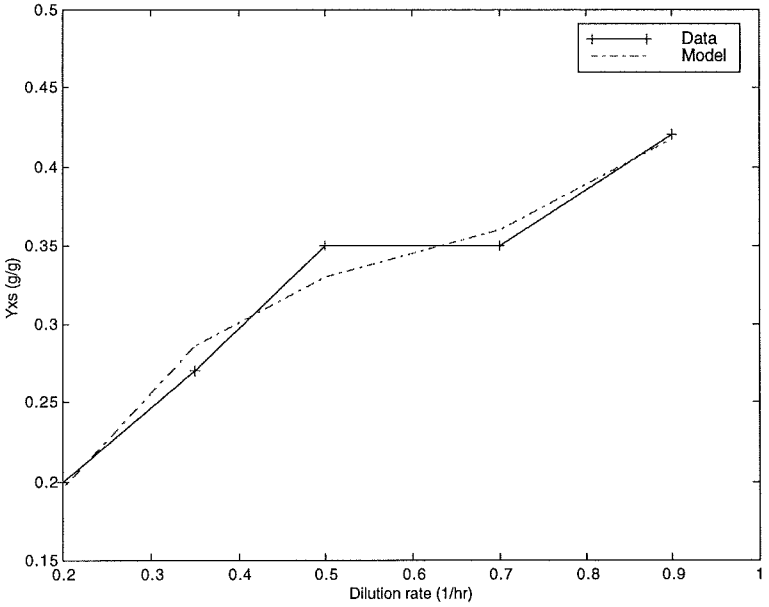


Figure 15. The least-squares fit for the yield $Y_{X/S}$

The result of the identification of the parameters λ and ϕ in eq. (5) for the specific growth rate μ is shown in Figure 16(top).At each sample, μ is put equal to the measured dilution rate. The identification gave a good estimation of the specific growth rate. The identification of the parameter m in the term rS in eq. (7) was more difficult. The result of the identification is shown in Figure 16(bottom).

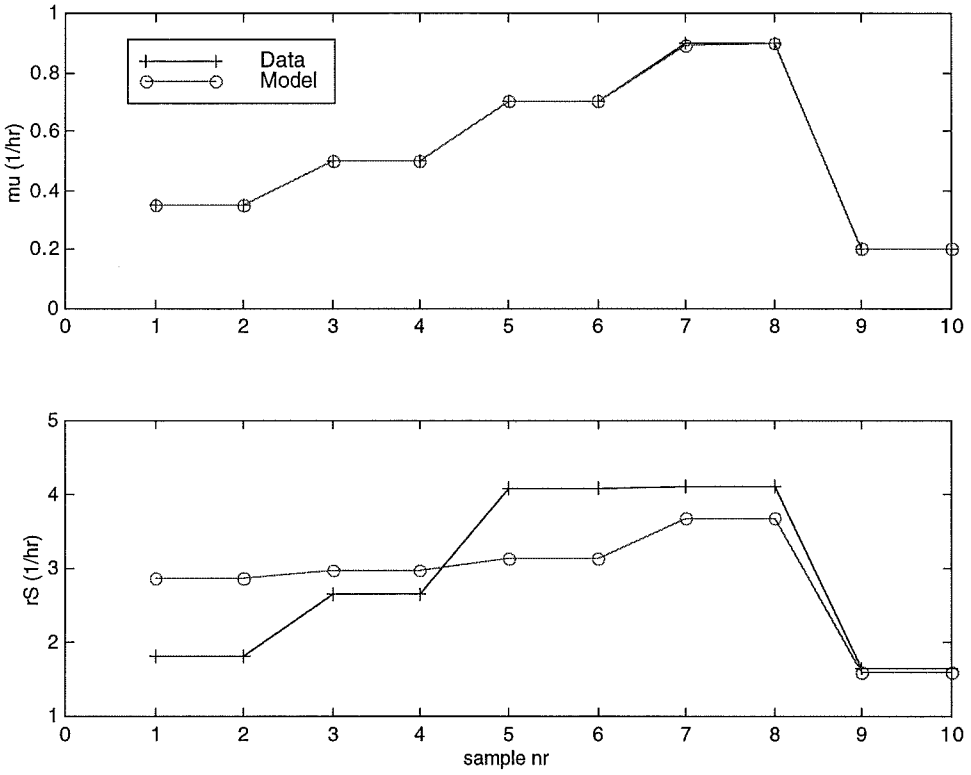


Figure 16(top). The least-squares fit for the specific growth rate μ **(bottom)**. The least-squares fit for the term rS

The reason why the least-squares fit did not work properly was probably that the data points acquired were to few. The measurement errors of the biomass concentration can be seen in Figure 17 where the acquired data and the model are compared. The measurement error is bigger for the biomass than for the substrate and product concentrations since the evaluation of the samples contains more steps for the biomass than for the substrate and product concentrations. The poor fit to the data made that there are steady-state errors for the lactose concentration. But since the lactose concentration was not used in the linearization of the model, the identification was left like this. The reason why the biggest error is for the lowest dilution rate is that at lower dilution rates D, the term rS has higher influence on the substrate consumption according to eq. (7). The parameters are given in Table 6.

Identified parameters		Parameters provided by Nestlé	
λ	0.0796	μmax	0.99 h^{-1}
ϕ	0.1529	α	2 g/g
m	1.1264 h^{-1}	β	1.4 h^{-1}
$Y_{X/S}(D)$	$1.3339D^3-2.4621D^2+1.6519D-0.0471$ g/g		

Table 6. The identified and provided parameters in the fermentation model

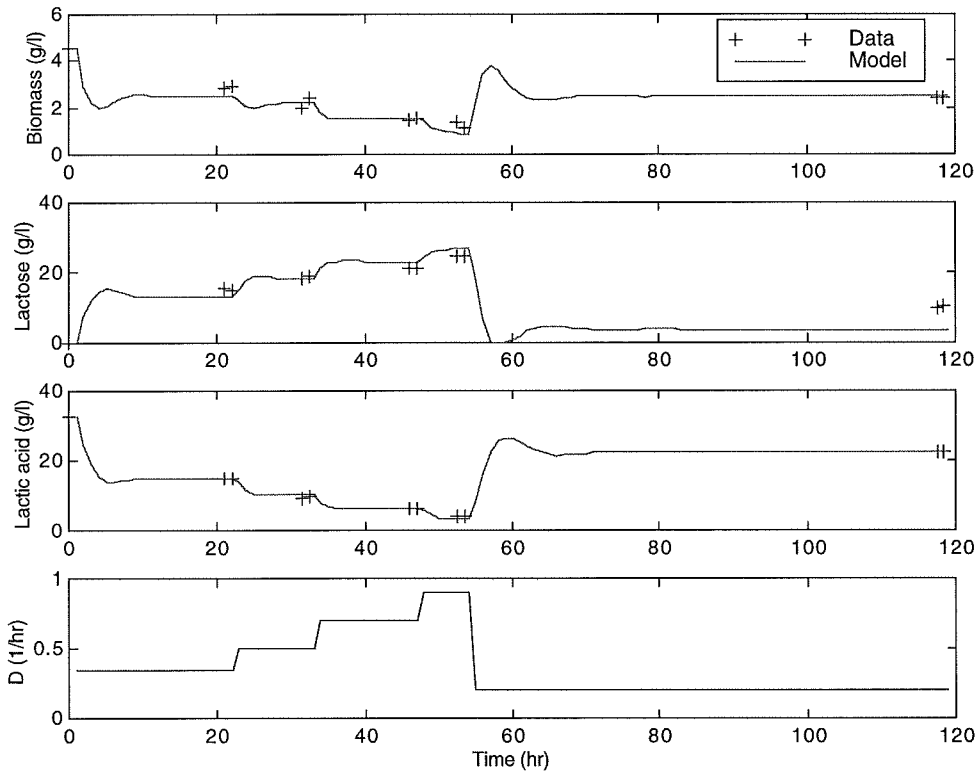


Figure 17. Biomass, Lactose and Lactic acid verses time.

The validation of the model is shown in Figure 18 during two stepchanges in the dilution rate. Here the samples are taken during the transient start-up times. The stepchanges were altered or ended before the system reached steady-state but the dynamics of the fermentation was modeled satisfactory.

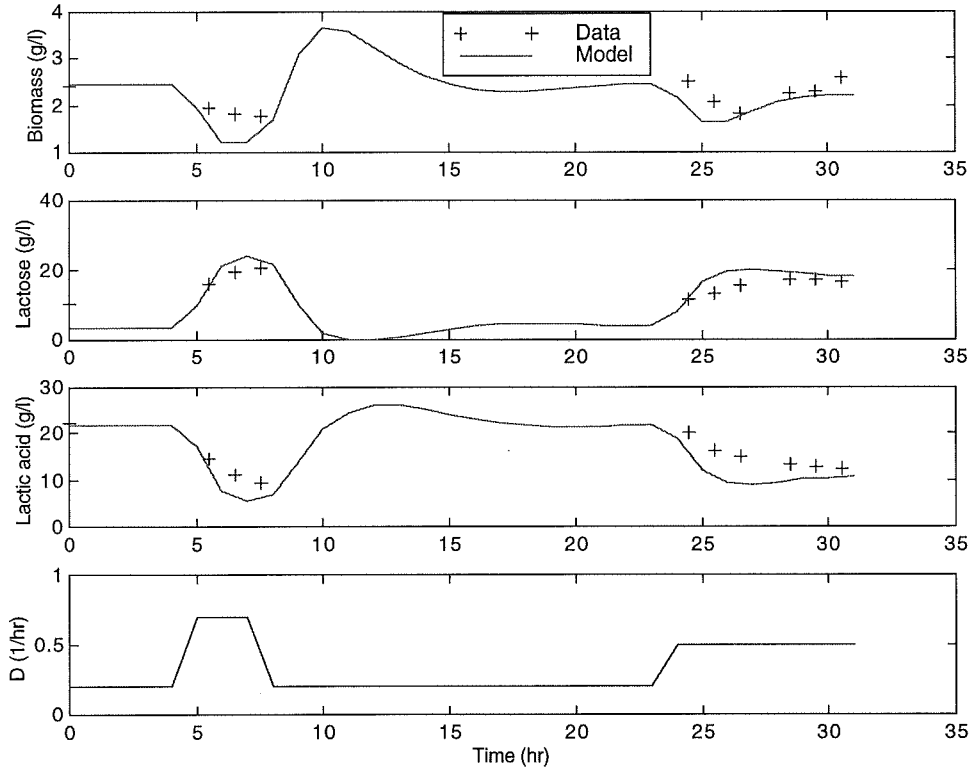


Figure 18. Validation of the fermentation model on transients

7.1.2 The cell recycle unit

The output from an identification attempt with a PRBS signal is shown in Figure 19. The plot shows that the filter has some initial dynamics. In the beginning, the output is drifting downwards and stabilizes after about 10 minutes. The explanation for this behaviour is that when the stable flow is reached, the drag force acting on the fouling particles (the cells) is balanced by the lift force generated by the shear field near the filter membrane. Both the final stable flux, as well as the time required to reach it are functions of the fluid being filtered. In Figure 20, a PRBS run is shown with the data acquisition started after 10 minutes. Since the data now had stabilized, an identification was done on with the data set detrended around zero to get rid of the output offset.

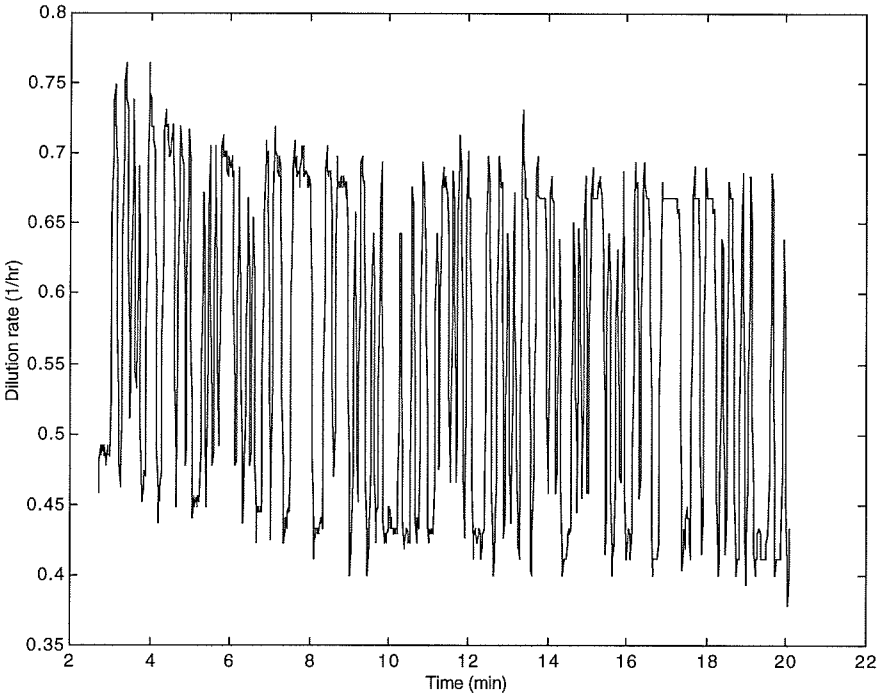


Figure 19. Output from a PRBS input signal with data acquisition started after 3 minutes

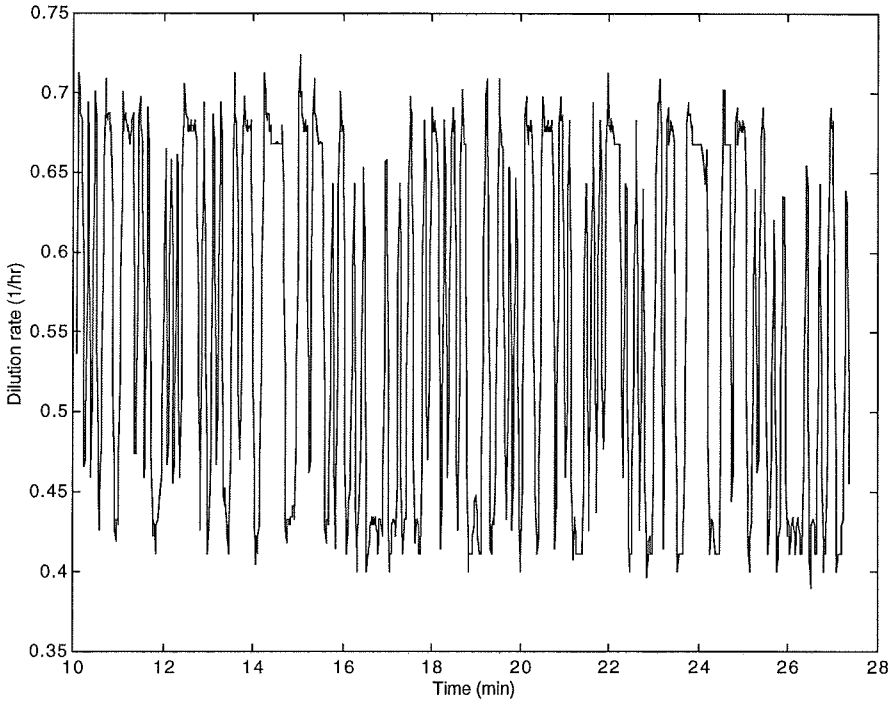


Figure 20. Output from a PRBS input signal with data acquisition started after 10 minutes

The result of the identification procedure described in *Appendix 2* was that the ARMAX model in eq (11) gave a smaller loss function than the ARX model in eq (10). The number of parameters of the identified polynomials A, B and C are called n_A, n_B and n_C respectively and the delay is called n_k. The identified model structure was

n_A	3
n_B	1
n_C	3
n_k	3

Table 7. The chosen number of parameters of the polynomials A,B,C and the delay n_k

The identified ARMAX model was

$$\begin{aligned} A(q^{-1}) &= 1 - 1.7188q^{-1} + 1.1570q^{-2} - 0.2890q^{-3} \\ B(q^{-1}) &= 0.012478 \\ C(q^{-1}) &= 1 - 0.8989q^{-1} + 0.2359q^{-2} + 0.0591q^{-3} \end{aligned} \tag{27}$$

But as mentioned above, the model depends on which state the filter is in and on the fluid being filtered. The model can thus change during the cell recycle. The B polynomial is only a constant since the delay operator q^{-1} is included in the delay q^{-nk}. In Figure 21 the pole-zero plot of the identified model of the cell recycle unit is shown.

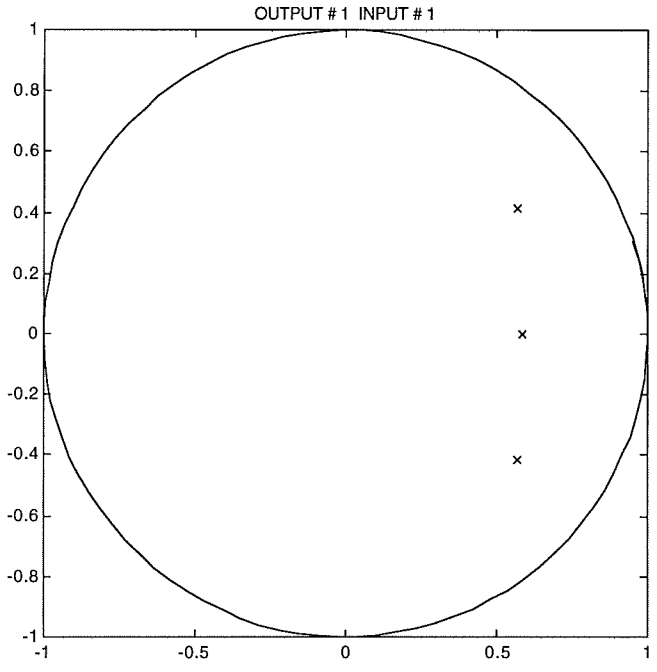


Figure 21. The pole-zero plot of the identified cell recycle unit model

To validate the model, step changes in the input signal are shown in Figure 22. The delay between the input signal and the output signal and the steady-states were modeled satisfactory. The disturbances on the measured signal can be due to air-bubbles and uncertainty in the balance readings. An estimation of the time constant τ_m for the process to reach 66% of the stepchange gave τ_m=*13 seconds.*

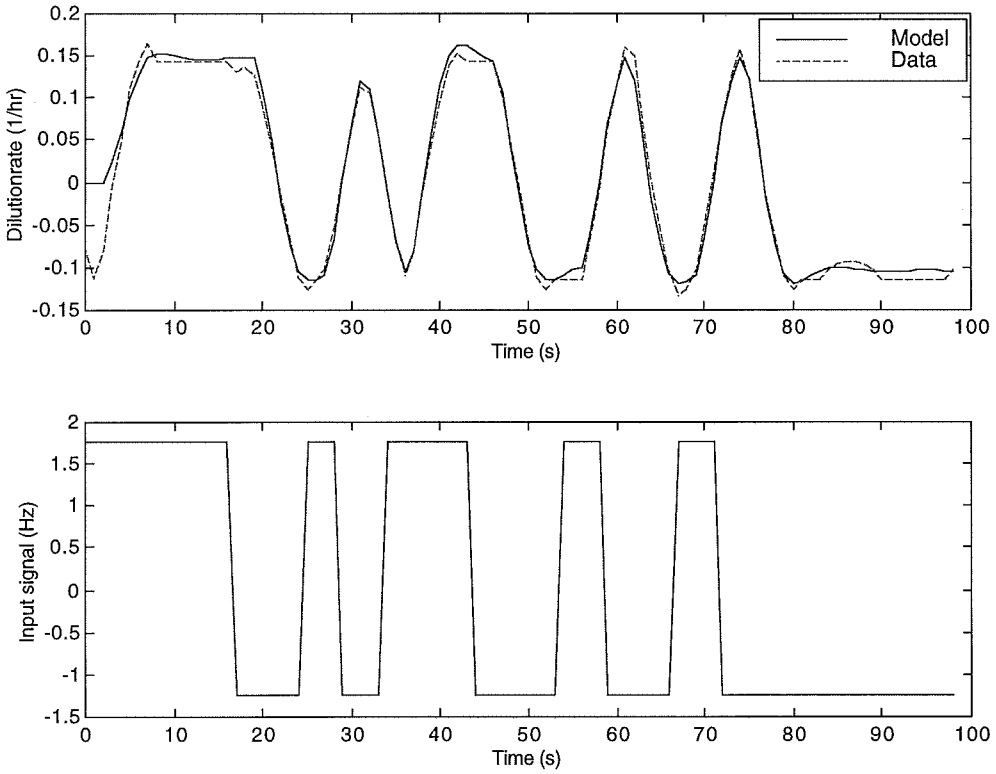


Figure 22. Measured and estimated output during stepchanges in the input signal for the cell recycle unit.

7.2 Simulations

A simulation for the whole system containing both the CSTF and cell recycle model is shown in Figure 23. The setpoint for the product concentration, 18 g/l, and the other conditions for the system are the same as in chapter 5.1. The cascade control was simulated with a sample time of two minutes. The inner control loop was represented by the closed loop transfer function of the cell recycle unit. Measurement noise of amplitude 0.1 was added to the simulation. The system started with a full cell recycle and when the product concentration had reached the operating set point, the fermenter controller was started. The result was that the biomass had been concentrated to ca 10 g/l.

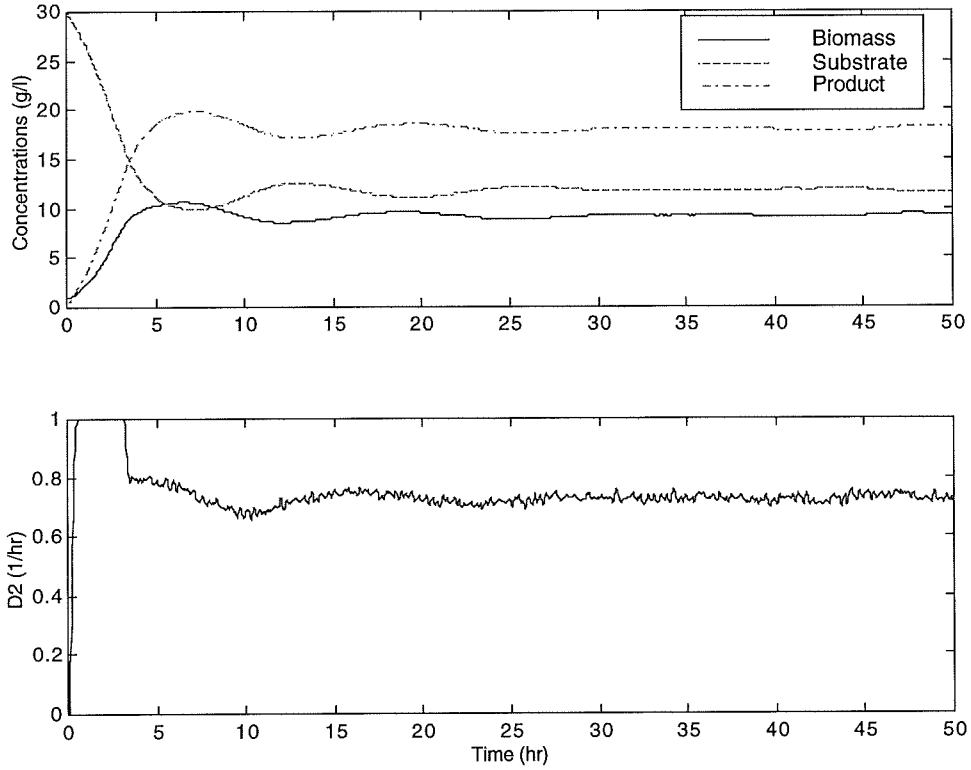


Figure 23. A simulation of the whole system containing both the CSTF and cell recycle model

The simulation in Figure 24 was made to evaluate how high cell concentration that could be achieved of the system. The overall dilution rate D was restricted to 1 hr^{-1} since no data for the yield $Y_{X/S}$ was provided for higher dilution rates. The desired biomass concentration was set to 13 g/l and the filter efficiencies were all put to 0.9 (1 means ideal filtration, 0 no filtration). The plot shows that the controller manage to control the biomass concentration to the desired setpoint and that the dilution rate D_2 can be increased further to achieve higher biomass concentrations. But there is a trade off between how high biomass concentration which can be achieved in the fermenter and how much biomass which can be extracted from the system. If the aim is only to maximize the biomass concentration, a total cell recycle can be done. But then the system never reaches steady-state and no biomass is extracted from the system. Therefore a bleed dilution rate D_1 is needed to be able to control the system and to have a production of biomass.

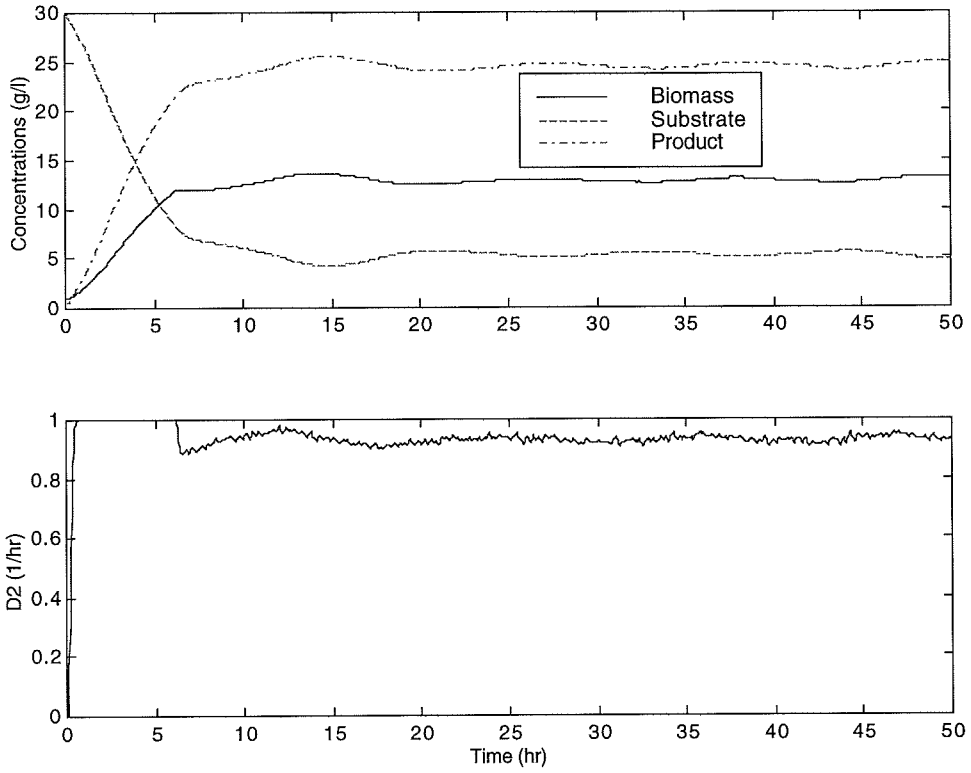


Figure (24). A simulation of the CSTF with the cell recycle unit. The desired biomass concentration was 13 g/l

The idea was that the product concentration would be measured online via calculations on the base added. But the measured differences between the base added at different product concentrations were to small or made no sense. The reasons for this could be that the medium flow was not "neutralized" before added to the fermenter, the hysteresis in the pH controller could be too big and that the base added was to strong. The pH fluctuated between 5.46 and 5.52.

Conclusions

The simulations show that it is possible to concentrate the cells with cell recycling to 10-13 g/l at an overall dilution rate D of 1 h^{-1}. The identification of the CSTF model could probably be more accurate with more data points. Since the model of the cell recycle unit is changing with the filter performance, an online estimation of the model is preferred. The controller for the cell recycle unit could then be an adaptive controller which is more suitable with a changing process.

Acknowledgements

I wish to thank my supervisors Mr. Sergio Valentinotti and Dr. Ulf Holmberg for their valuable assistance to this master's thesis and likewise Prof. Dominique Bonvin for his counsel. At Nestlé I want to thank Ms. Anne Bouché and Mr. Pierre-Alain Richon for helping me in the laboratory and also Dr. Peter Niederberger for his advice. My gratitude is also sent to l'Institut d'Automatique for the friendly atmosphere there.

8 References

1. Lee J.M.(1992), *Biochemical Engineering,* Prentice-Hall International Series in the Physical and Chemical Engineering Sciences

2. Taniguchi M., Kotani N., and Kobayshi T. (1987), "High-Concentration Cultivation of Lactic Acid Bacteria in Fermentor with Cross-Flow Filtration", *Journal of Fermentation Technology*, 65(2), 179-184

3. Ljung L. (1987), *System Identification : Theory for the user,* Prentice-Hall Information and System Sciences Series

4. Åström K.J. and Wittenmark B. (1984), *Computer Controlled Systems*, Prentice-Hall Information and System Sciences Series

5. Longchamp R. (1995), *Commande Numerique de Systèmes Dynamiques,* Presses Poly-techniques et Universitaires Romandes, Lausanne

9 Appendix

9.1 General least-square

In the general least-square problem the model value $\hat{y}(t)$ of the measured value *y(t)* can be written as

$$\begin{aligned} &\hat{y}(t) = \varphi_1(t)\theta_1 + \varphi_2(t)\theta_2 + \cdots + \varphi_n(t)\theta_n = \varphi(t)^T\theta \\ &\theta^T = \begin{bmatrix}\theta_1 & \theta_2 & \cdots & \theta_n\end{bmatrix} \\ &\varphi(t)^T = \begin{bmatrix}\varphi_1(t) & \varphi_2(t) & \cdots & \varphi_n(t)\end{bmatrix} \end{aligned} \qquad \text{(A1)}$$

where θ_i are the unknown parameters, and φ_i are known functions. The variables φ_i are called *the regression variables* or the *regressors*. Since the model value $\hat{y}(t)$ is linear in the parameters and the least square criteria (9) is quadratic, the problem of minimizing the loss function admits an analytical solution.

Introduce the notation

$$Y(t) = \begin{bmatrix} y(1) & y(2) & \cdots & y(t) \end{bmatrix}^T$$

$$\Phi(t) = \begin{bmatrix} \varphi^T(1) \\ \vdots \\ \varphi^T(t) \end{bmatrix}$$

If the matrix $\Phi^T\Phi$ is nonsingular the minimum is unique and given by

$$\hat{\theta} = \left(\Phi^T\Phi\right)^{-1}\Phi^T Y \qquad \text{(A2)}$$

Usually the regression variables are the input signals *u(t)* and output signal *y(t)*. If the input signal contains sufficiently many frequencies, the matrix $\Phi^T\Phi$ is nonsingular in eq. (A2) and a least-square estimate of the parameters in the *A* and *B* polynomials can be done. For more details, the reader is referred to any digital control book, for example Åström (4).

9.2 Identification procedure for the cell recycle unit

To determine the global order $N=max(n_A, n_B+n_k-1)$, the models in eq (10) and (11) were identified with the order $n=n_A=n_B=n_C$ with n in the interval of 1 to 10. The delay n_k was set to 1 which means that n equals the global order N. The estimated variance of ε, which is called the loss function V verses the global order N is shown in Figure A1.

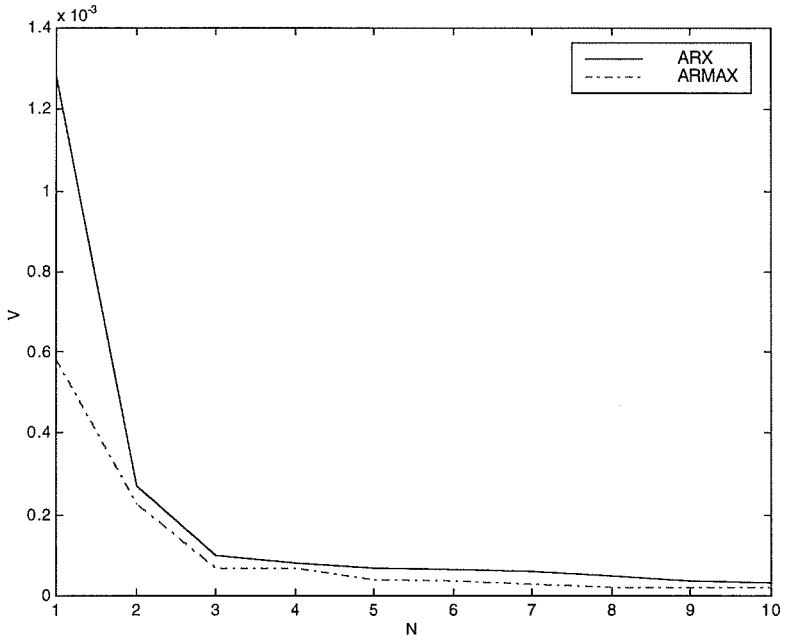


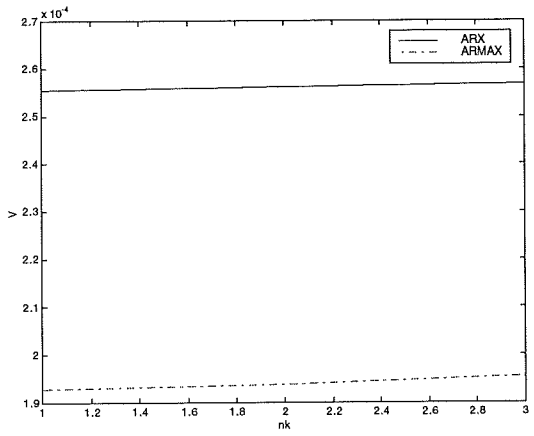
Figure A1. Loss function V=variance of ε verses global order $N=max(n_A, n_B+n_k-1)$

From Figure A1 it can bee seen that increasing the global order above $N=3$ does not lower the loss function V much. The global order was therefore chosen as $N=3$. To be able to determine the structure of the model, one of the orders or the delay was changed and the others were set according to the global order.

To determine the delay n_k, n_A was fixed at $n_A=N$, $n_C=1$, and the models were identified with n_k in the interval of 1 to N. The loss function V is shown in Figure A2 (a). For the ARMAX model there is no big difference but to minimize the loss function for the ARX model according to Figure A2 (a), $n_k=1$ should be chosen. But a delay of one led to that the first coefficients in the B polynomial became small compared to the coefficients in the A polynomial. That means if a small change is done in the A polynomial, the parameter estimation of the B polynomial becomes useless. Therefore the delay was increased to $n_k=3$ which increased the coefficients in the B polynomial. Furthermore, the control signal was divided with a factor 100 during the identification in order to reduce the difference between the coefficients in the A and B polynomials. Before presented in (27), the identified B polynomial was divided with a factor 100 to compensate that.

A look at the pole-zero plot of the model can also be useful to determine the delay. There is rule of thumb that says that if there are unstable zeros, the delay could be under-estimated. On the other hand, if all the zeros are stable the delay could be over-estimated. In Figure A2 (b) the pole-zero plot for the ARMAX model when $n_k=3$ is shown. Neglecting the zero in infinity, all zeros are stable and thus indicating that the delay could be decreased. When the delay was decreased, a zero became unstable indicating that the delay could be increased. But as mentioned above, the delay was chosen $n_k=3$.

(a)



(b)

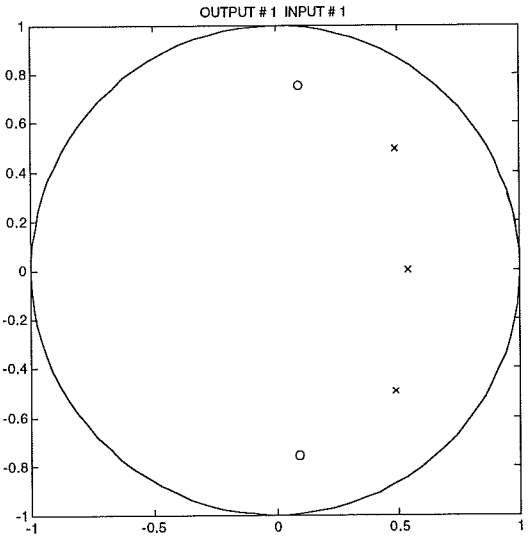
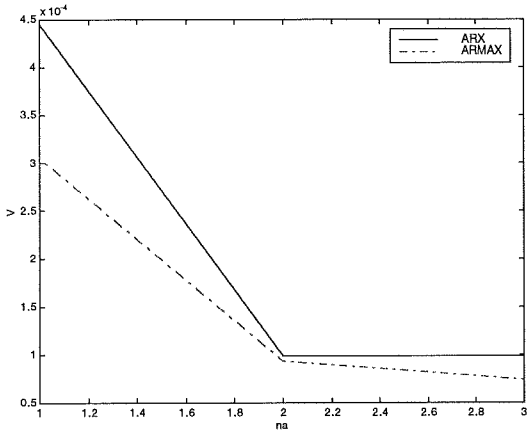


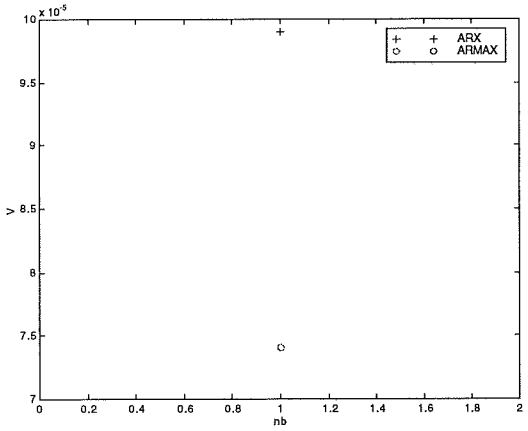
Figure A2(a). Loss function V verses the delay n_k **(b)** Pole-zero plot for ARMAX $n_k=3$

To determine the order of the A polynomial, n_A, n_k was fixed at the optimal vale 3, $n_C=1$, and the models were identified with n_A in the interval of 1 to N. The loss function V is shown in Figure A3(a). The value of n_A was set to 3 in order to minimize the loss function.

(a)



(b)



(c)

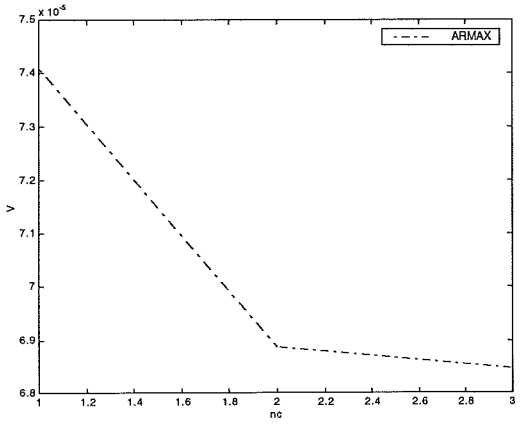


Figure A3(a). Loss function V verses the order of the A polynomial n_A,(a), **(b).** Loss function V verses the order of the B polynomial $n_B=1$, **(c).** Loss function V verses the order of the C polynomial n_C

Since the order of the B polynomial $n_B=N-n_k+1=3-3+1=1$, there existed no more alternatives than $n_B=1$. The loss function is shown in Figure A3(b). The ARMAX model was chosen since it gave a smaller loss function value. The loss function verses the order of C polynomial n_C, can be seen in Figure A3(c). It was set to 3 in order to minimize the loss function. The chosen orders of the polynomials and the delay are given in Table A1.

n_A	3
n_B	1
n_C	3
n_k	3

Table A1. The chosen orders of the polynomials A,B,C and the delay n_k