



Versatile modeling and optimization of fed batch processes for the production of secreted heterologous proteins with *Pichia pastoris*

Michael Maurer and Brigitte Gasser, IAM / DBT
Manfred Kühleitner, Institut für Mathematik / DIB
Diethard Mattanovich, School of Bioengineering / FH Campus Vienna

While the mathematical description of processes like growth and product formation have been fairly well achieved, it is still not routine practice to model and optimize fed-batch processes by mathematical means. A typical case of fed-batch process is the production of recombinant proteins with microorganisms or mammalian cells. As the production of many proteins in yeasts is quite cost sensitive, it will be highly desirable to have a tool available that allows a simple yet reliable prediction of productivity, process time and product titers. We have developed a mathematical model which can be used in MS Excel™ with the Solver™ software to optimize the time course of the media feed in order to maximize the volumetric productivity or Space Time Yield. The advantages of the procedure we describe here are the ease of use and the flexibility applying software familiar to every scientist and engineer and rapid calculation which makes predictions extremely easy, so that many options can be tested *in silicio* quickly.

1. Method of Calculation. We divide the total feed period in equal intervals $i_n=[t_n, t_{n+1}]$ for $(1 \leq n \leq N)$ of length dt . We start with an initial value $dt=1h$. The best value for dt is determined within the optimization process.

At every point of time t_n we denote by $X_n=X(t_n)$ the amount of biomass and by $P_n=P(t_n)$ the amount of product in the bioreactor. At the beginning of the fed-batch process the initial values are $X(0)=X_0$ and $P(0)=P_0$.

First we have to describe the growth of the biomass. We use the simplest model, the exponential growth model,

$$(1) \quad \frac{dX}{dt} = \mu \cdot X$$

Since the specific growth rate μ of the biomass depends on time, we calculate the above equation in discrete time steps

$$(2) \quad X_{n+1} = X_n \cdot e^{\mu_n \cdot dt}$$

where μ_n is the specific growth rate during the interval i_n . The initial values for μ_n are chosen arbitrary, for instance $\mu_n = \mu_{max}$. The optimal values for all the μ_n 's are determined within the optimization process.

Second we have to describe the accumulation of the product. We simply calculate the total product yield during the interval i_n by the following formula

$$(3) \quad P_{n+1} = P_n + dP_n \text{ with } dP_n = q \cdot X_n \cdot dt.$$

The relationship between the specific rate q of product formation and the specific growth rate μ was experimentally determined in chemostat cultures. The dependence of q on μ was described analogous to Monod equation:

$$(4) \quad q = q_{max} \frac{\mu}{k + \mu}$$

The values for q_{max} and k are derived from the experimental data by the method of least squares, i.e. the parameters q_{max} and k are chosen that the sum of the deviations from the experimental data squared is minimal.

Next we have to calculate the amount of substrate dS which we must feed in the time interval i_n . To do this, let S_n be the amount of substrate added to the bioreactor until the time point t_n . Then the substrate consumption rate depends on the amount and on the increase of biomass, i.e.

$$(5) \quad \frac{dS}{dt} = -\left(\frac{\mu}{Y_{XS}} + m\right) \cdot X \cdot dt$$

where m is the maintenance coefficient and Y_{XS} is the true yield coefficient of biomass from substrate. Inserting formula (1) in (5) the amount of substrate feed in the interval i_n calculates as

$$(6) \quad dS_n = \left(\frac{\mu_n}{Y_{XS}} + m\right) \cdot X_n \cdot dt_n$$

To calculate the parameters Y_{XS} and m from experimental data of chemostat cultures by the method of least squares, we use the observed biomass yield coefficient Y'_{XS} depending on the specific growth rate μ .

This is done by $dX = -Y'_{XS} \cdot dS$ and inserting formula (1) and the formula for the whole substrate consumption which implies

$$(7) \quad Y'_{XS} = \frac{\mu}{\mu/Y_{XS} + m}$$

Last but not least we need the total volume for the calculation of the volumetric productivity. The model process starts with a batch volume of $V=1L$. The total volume at each time interval is then

$$(8) \quad V_{n+1} = V_n + \frac{dS_n}{1000 \cdot s \cdot \rho}$$

with the substrate concentration in the feed medium s and the density of the feed medium ρ . Due to high biomass concentrations achieved in *P. pastoris* fermentation, the cells occupy a significant fraction of the total volume, while the product is secreted to the liquid phase, the culture supernatant. In order to calculate the product concentration, the available liquid volume V_l is calculated at each time interval with the specific volume of wet biomass, which is derived from dry biomass as the specific volume per dry biomass $v_{YDM}=0.0033 \text{ L g}^{-1}$.

$$(9) \quad V_n = V_n - X_n \cdot v_{YDM}$$

Finally we calculate the biomass and the product concentrations x and p . They calculate at the time point t_n by the following formulas

$$(10) \quad x_n = \frac{X_n}{V_n} \text{ resp. } p_n = \frac{P_n}{V_n}$$

The medium feed rate F_n at each time point is

$$(11) \quad F_n = \frac{dS_n}{s} \frac{1}{dt}$$

2. Optimization. The goal of our optimization problem is to find the best values for the specific growth rates μ_n and the best value for dt such that the volumetric productivity Q calculated at the time point t_{N+1} as

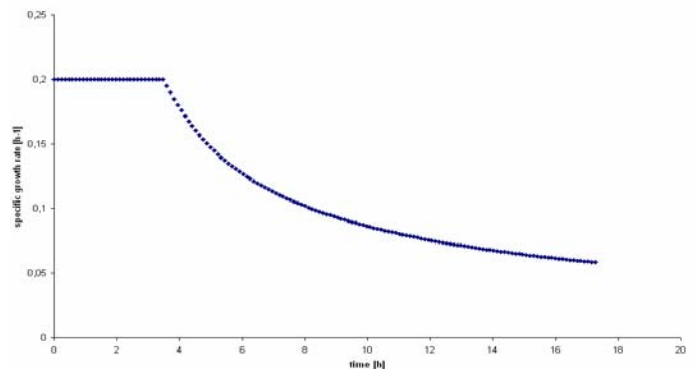
$$(12) \quad Q_{N+1} = \frac{P_n}{t_{N+1} \cdot V_{N+1}}$$

is maximized under the following constraints:

$$(13) \quad \mu_{min} \leq \mu_n \leq \mu_{max} \text{ for } (1 \leq n \leq N) \text{ and } X(t_{N+1}) = X_{max}$$

Here $\mu_{max}=0,2h^{-1}$ is the maximum specific growth rate at just below washout in chemostat cultures. Since below $\mu=0,02h^{-1}$ significant product degradation appeared, the lower boundary was set at $\mu_{min}=0,03h^{-1}$. Also the biomass concentration need to be limited. The upper limit is mainly defined by the cell separation step, which is practically limited with approximately $100 \text{ g} \cdot \text{L}^{-1}$ dry mass.

Optimized time course of specific growth rate



The optimal feed phase consists of a constant specific growth rate at the beginning followed by a phase with steadily decreasing specific growth rate.