Application of Genetic Algorithms to Model Parameter Identification of a Recombinant *E. coli* High-Cell Density Fed-batch Fermentation

Isabel Rocha and Eugénio C. Ferreira

Centro de Engenharia Biológica, Universidade do Minho 4710 - 057 Braga PORTUGAL
e-mail: {irocha, ecferreira}@deb.uminho.pt

Abstract

In this work, a genetic algorithm was used to estimate both yield and kinetic coefficients of an unstructured model representing a fed-batch high-cell density fermentation of *E. coli*.

The model, based on the General State Space Dynamical Model was used to represent the three major metabolic pathways: oxidative growth on glucose, fermentative growth on glucose, and oxidative growth on acetate. The structure of the kinetic equations was derived from literature and adapted to represent experimental results.

Genetic Algorithms were used to minimize the normalized quadratic differences between simulated and real values of the state variables X, A and W, by manipulating both yield and kinetic coefficients.

I. INTRODUCTION

Fed-batch fermentation of *E. coli* can be modelled by the following balance equations:

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ A \\ W \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -k_1 & -k_2 & 0 \\ 0 & k_3 & -k_4 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} X - D\begin{bmatrix} X \\ S \\ A \\ W \end{bmatrix} + D\begin{bmatrix} 0 \\ S_{in} \\ 0 \\ W \end{bmatrix}$$

With X, S and A are biomass, glucose and acetate concentrations; W is the culture weight; D is the dilution rate μ_i are the specific growth rates and k_i are yield coefficients.

Parameter estimation can be seen as an optimization problem where the performance criterion is the sum of squared errors between measured and simulated data:

$$J = \frac{\sum_{i=1}^{t}\left[\frac{X_{sim,i}-X_{exp,i}}{X_{exp,aver}}\right]^2}{\sigma_X^2} + \frac{\sum_{i=1}^{t}\left[\frac{A_{sim,i}-A_{exp,i}}{A_{exp,aver}}\right]^2}{\sigma_A^2} + \frac{\sum_{i=1}^{t}\left[\frac{W_{sim,i}-W_{exp,i}}{W_{exp,aver}}\right]^2}{\sigma_W^2}$$

II. METHODS

The experimental setup used to run the fed-batch fermentations with its main components - fermenter, digital control unit, balances, pumps, mass spectrometer, on-line filtration device and FIA system is illustrated in Figure 1.
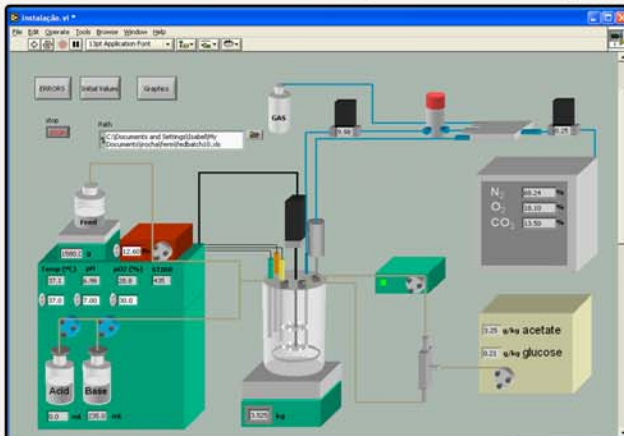

Fig. 1 Control panel of the supervisory program for acquisition and control of the fed-batch fermentation of *E. coli* representing the main equipment and devices.

Simulation of the fermentation was conducted using MATLAB version 6.0 using a Runge-Kutta 4th order integration with fixed stepsize.

The Genetic and Evolutionary Algorithm Toolbox (GEAtbx 3.3) for MATLAB developed by Pohleim was used for optimization purposes.

III. RESULTS

During optimization, the model described was used to simulate fed-batch fermentations by applying the same feeding profile as in real fermentations. Then, the state variables X, A and W were compared with the real ones using the performance criterion. Genetic algorithms were used to generate an initial population of yield and kinetic coefficients for the simulation and to obtain iteratively an optimum set of those parameters.

The optimization converged to an optimum after 1000 iterations and the corresponding results from a validation of the identified coefficients are shown in Fig. 2.
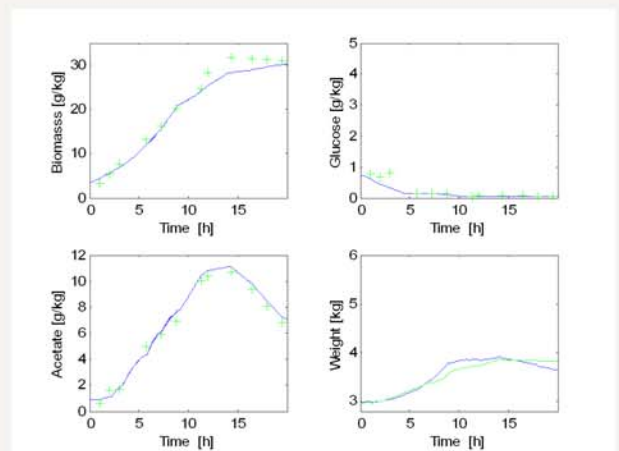

Fig. 2 Validation of a parameter set in coefficients identification. Simulated variables using the estimated parameter set are in blue color, while measured variables are shown in green. The symbol (+) denotes off-line measurements.

IV. CONCLUSIONS

Using this methodology, it is possible to estimate with great accuracy several model parameters, without needing extensive mathematical manipulations of the model. As the results show, there is a very good agreement between real and simulated values for all measured state variables.