redin Revista Facultad de Ingeniería, Universidad de Antioquia, No.87, pp. 35-40, 2018

Multi-objective optimization in biotechnological processes: application to plant cell suspension cultures of *Thevetia peruviana*

Optimización multi-objetivo en procesos biotecnológicos: aplicación al cultivo de células vegetales en suspensión de *Thevetia peruviana*

Adriana Patricia Villegas-Quiceño[1*], Juan Pablo Arias-Echeverri[2], Daira Aragón-Mena[4], Silvia Ochoa-Cáceres[3], Mario Evelio Arias-Zabala[2]

[1]Grupo Termomec, Facultad de Ingeniería, Universidad Cooperativa de Colombia. Calle 50 # 41-34. Bloque A piso 4. A. A. 050012. Medellín, Colombia
[2]Grupo de Biotecnología Industrial, Facultad de Ciencias, Universidad Nacional de Colombia Sede Medellín. Calle 59A # 63-20. A. A. 050034. Medellín Colombia
[3]Grupo de Investigación en Simulación, Diseño, Control y Optimización de Procesos (SIDCOP), Facultad de Ingeniería, Universidad de Antioquia. Calle 67 # 53 – 108. A. A. 050010. Medellín Colombia
[4]Audubon Sugar Institute, Louisiana State University Agricultural Center. 3845 Highway 75. C. P. 70776. St Gabriel, Louisiana, United States.

ARTICLE INFO:
Received December 12, 2017
Accepted April 13, 2018

KEYWORDS:
Model-based optimization, Modelling and optimization, Multi-objective problem, Mechanistic models, Biotechnological processes

Optimización basada en modelos, Modelado y optimización, Problema multiobjetivo, Modelos mecanísticos, Procesos biotecnológicos

ABSTRACT: Bioprocesses productivity is a compromise between two conflicting objectives, maximization of biomass growth rate and minimization of substrate consumption. In this work, a model based multi-objective optimization problem is solved for improving the process productivity in plant cell suspension cultures of *Thevetia peruviana*. A solution of the multi-objective problem allowed determining the optimal initial concentrations of substrate and biomass for assuring maximal productivity. Model-based optimization is carried out using a mechanistic model, which includes a representation of the intracellular processes taking place on the plant cells. The best solutions were chosen from the Pareto front in agreement with expert criterion. Results indicate that an initial inoculum concentration of 3.91g/L and an initial sucrose concentration of 23.63g/L, are recommended as initial conditions for obtaining a biomass productivity of 1.57g/L*day with an acceptable sucrose uptake. Experimental validation of the optimal found was carried out and the productivity obtained was 1.52g/L using an initial inoculum concentration of 4.27g/L and an initial sucrose concentration of 25.44g/L. Results suggest that the proposed methodology can be extended to increase the productivity in terms of metabolite production from this plant cell cultures and other plant species.

RESUMEN: La productividad de los bioprocesos es un compromiso entre dos objetivos en conflicto, la maximización de la velocidad de crecimiento de la biomasa y la minimización del consumo de sustrato. En este trabajo, se resuelve un problema de optimización multi-objetivo para mejorar la productividad del cultivo en suspensión de células vegetales de la especie *Thevetia peruviana*. La solución del problema multi-objetivo permitió determinar las concentraciones iniciales óptimas de sustrato y biomasa para garantizar la máxima productividad. La optimización se lleva a cabo utilizando un modelo mecanistico, que incluye una representación de los procesos intracelulares que tienen lugar en las células vegetales. Las mejores soluciones se eligieron del frente de Pareto teniendo en cuenta el criterio experto. Los resultados indican que se recomienda una concentración inicial de inóculo de 3.91g/L y una concentración inicial de sacarosa de 23.63g/L como condiciones iniciales para obtener una productividad de biomasa de 1.57g/L*día con un consumo aceptable de sacarosa. Se llevó a cabo la validación experimental del óptimo encontrado y la productividad obtenida fue de 1.52g/L usando una concentración de inóculo inicial de 4.27g/L y una concentración inicial de sacarosa de 25.44g/L.

* Corresponding author: Adriana Patricia Villegas Quiceño
E-mail: adriana.villegas@ucc.edu.co
ISSN 0120-6230
e-ISSN 2422-2844

DOI: 10.17533/udea.redin.n87a05

Los resultados sugieren que la metodología propuesta puede ampliarse para aumentar la productividad en términos de producción de metabolitos a partir de estos cultivos de células vegetales y otras especies vegetales.

1. Introduction

Model-based optimization can be used in biotechnological processes to design biological systems, to enhance the metabolite production and biomass production, and to ensure the optimal manipulation of the genetic and metabolic composition among others [1]. On the other hand, in metabolic transformation by cellular systems, many criteria require to be satisfied simultaneously, for instance, to maximize metabolic reaction rates and steady state fluxes and to minimize concentrations of metabolites in transient times. In general, metabolic variables (enzyme, substrates and metabolic concentrations) should remain within a certain physiological range in order to avoid undesirable effects and maintain cell viability [2, 3]. Along these lines, multi-objective optimization finds solutions that are optimal for several conflicting objectives simultaneously [4]. These solutions are called Pareto optimal solutions. Despite the existence of multiple Pareto solutions, in practice, only one of these solutions can be chosen, generally based on previous knowledge [5]. Multi-objective optimization has been reported [4] to maximize ethanol concentration and minimize each internal metabolite concentration for ethanol production from *Saccharomyces cerevisiae*. [6] used multi-objective optimization to estimate the kinetic model parameters of batch and fed batch fermentation processes for ethanol production using *Saccharomyces diastaticul*. [7] introduced a novel framework for the optimal development of biotechnological processes using optimization tools such as multi-objective mixed-integer nonlinear programming. Finally, [3] presented a methodology to maximize productivity in biotechnological processes using multi-objective optimization. The versatility of the methodology was demonstrated using a large-scale metabolic model of Chinese Hamster Ovary cells (CHO) [3].

Despite the significant benefits of multi-objective optimization, only a few metabolic engineering applications are found in the literature [3] and, to the author's knowledge, there are no available reports addressing the development of model-based multi-objective optimization using structured models for applications in plant cell cultures.

In this work, a multi-objective optimization to determine the initial extracellular sucrose concentration and initial inoculum concentration is used in order to increase biomass productivity maintaining low levels of substrate uptake in plant cell suspension cultures of *Thevetia peruviana*. The model used in this work corresponds to the structured model for plant cell suspension cultures presented by [8] for this specie. Finding the optimal initial conditions to maximize biomass productivity is the first step towards ensuring metabolite production in this plant cell culture.

2. Methodology

Figure 1 presents the methodology followed in this work to formulate and solve the multi-objective optimization problem. This methodology includes five steps, which are presented below.

2.1 Define the model

It is recommended to use a first-principles-based semi-physical model. The kinetic model should be selected taking into account the level of detail in the description of the phenomenon analyzed. According to the general classification of the kinetic models presented by [9], kinetic models in biotechnological processes can be classified as unstructured and structured models. Unstructured models focus on describing substrate uptake, cell growth and metabolite production at extracellular level. On the contrary, structured models present with certain level of detail the intracellular processes taking place. On the other hand, when the model considers aspects such as, size of the cells, viability, among others, it is necessary to differentiate between segregate and non-segregated models. It is important to highlight that the selected model must have all parameters estimated and it must be validated.

2.2 Define the optimization problem

The general multi-objective optimization problem can be stated as presented in Equations (1)-(5)

$$Maximize\left\{h_1\left(X\right), h_2\left(X\right), \ldots, h_k\left(X\right)\right\} \quad (1)$$

$$Minimize\left\{h_{k+1}\left(X\right), h_{k+2}\left(X\right), \ldots, h_{k+n}\left(X\right)\right\} \quad (2)$$

Subject to

$$f\left(\frac{dX}{dt}, X, t, \theta\right) = 0 \quad (3)$$

$$X\left(t_0\right) = X_0 \quad (4)$$

$$X_l \leq X_0 \leq X_u \quad (5)$$

Where, $X = [x_1, x_2, \ldots, x_r]$ indicates the state variables vector (i.e. biomass, substrates and metabolites

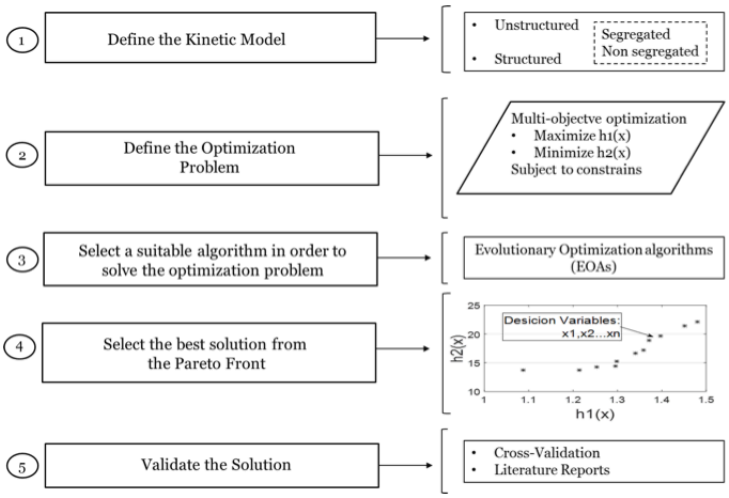

Figure 1 Methodology for multi-objective optimization. Based on the works by [3, 5, 10]

concentrations). f is the set of equality (i.e. algebraic or ordinary differential) constrains describing the system dynamics, and $\{h_1(X), h_2(X), \ldots, h_k(X)\}$ and $\{h_{k+1}(X), h_{k+2}(X), \ldots, h_{k+n}(X)\}$ are the conflicting optimization objectives to be optimized t is the time, $\theta = [\theta_1, \theta_2, \ldots, \theta_m]$ is the vector of model parameters and X_0 is the vector of initial conditions. Equation (5) corresponds to the boundaries of the decision variables.

2.3 Select a suitable algorithm in order to solve the optimization problem

Among the different approaches for solving multi-objective optimization problem, Evolutionary Multi-objective optimization (EMO) methods have emerged, and are the most popular [10]. EMO makes reference to the use of Multi-Objective Evolutionary Algorithms (MOEAs) such as: Genetic Algorithms (GAs), Evolution Strategies (ESs), and Evolutionary Programming (EP) [10]. MOEAs are based on natural evolution and the Darwinian concept of "Survival of the Fittest". This implies aspects relative to reproduction, random variation, competition, and selection of contending individuals within the population [10]. In general terms, Evolutionary Algorithms (EAs) use a population based approach in which more than one solution participates in an iteration and evolves a new population of solutions in each iteration [5, 10]. Some advantages obtained using the evolutionary algorithms are: (1) These algorithms do not require derivative information, (2) They are relatively simple to implement and (3) multiple trade-off solutions can be found in a single simulation run [5, 10].

2.4 Select the Best Solution from the Set of Pareto Solutions

Solution to the multi-objective optimization problem results in a series of Pareto solutions. In most cases, the Pareto optimal set contains more than one element because there exist different trade-off solutions to the problem which offer different compromises between the objectives. In practice, solving a multi-optimization problem often means that a human decision-maker is involved using expert knowledge of the system under study [11]. However, when the number of Pareto optimal solutions is high, it could be difficult to select the best solution given the large set of alternatives. In this case, the best solution can be determined using the concept of "Knee solutions" [12]. A "knee" is defined as a region in the Pareto optimal front, which is visually a convex bulge in the front, and often constitutes the optimum in trade-off solutions to the problem [13]. In this work, the concept of "knee" solution is used in conjunction with the expert knowledge of the authors on the growth of plant cell suspension cultures of *Thevetia peruviana* for solving the optimization problem presented in section 3.

2.5 Validate the solution

Two approaches are generally used in order to validate the solutions obtained. In the first place, it is recommended to carry out an experimental cross validation using the optimal conditions selected from the Pareto front, for comparing the results obtained in real experiment, from those obtained by the mathematical solution of the problem. In a second approach, results delivered by the multi-objective optimization are compared with reports of the literature for similar biotechnological process [14, 15].

3. Multi-objective optimization in plant cell suspension cultures of *T. peruviana*

In this section, application of the methodology described in Section 2 is carried out for the case of plant cell suspension cultures of *T. peruviana*

3.1 Model description

The model used in this work is a mechanistic model proposed by [8] for describing the cell growth, substrates uptake and the main metabolites involved in the central metabolism in plant cell suspension cultures of *T. peruviana*. This model comprises 28 metabolic species, 33 metabolic reactions, and 61 parameters. Two sets of experimental data were used for parameter identification. Furthermore, a different data set was used for model validation. Experimental data sets were obtained carrying out different kinetic studies at different initial sugars concentration and inoculum concentration. pH was adjusted between 6.5 and 7.5. Each experiment was conducted during 18 days. The extracellular sugars were determined by an HPLC (Agilent), coupled to refractive index detector using a coregel 87P column and water as mobile. The biomass concentration was determined using the dry weight method.

Although this model is used as a case study to present the optimization strategies to maximize metabolite production, such strategies can be generalized for application to models of different plant cell cultures and biotechnological processes. The metabolic pathway considered in the model development is summarized model in Figure 2.

3.2 Definition of the optimization problem

An important variable in order to get the best performance in biotechnological processes is the productivity. The productivity p_i is a measure of the economic viability of the process and it is defined as the ratio between the variation on the concentration of the desired product and the variation of the time, i.e. the derivative of the biomass concentration with respect to the time as presented in Equation (6).

$$p_i(t) = dx_i(t)/dt \tag{6}$$

On the other hand, substrate concentration should remain within a certain physiological range in order to avoid undesirable effects and maintain cell viability.

In this work, this objective is defined as "Levels of Substrate" *LS* see Equation (7).

$$LS(t) = \sum_{j=1}^{q} \left| x_j(t) - x_{j0} \right| \tag{7}$$

Where $j = 1, 2, \ldots, q$ corresponds to the number of substrates involved, $x_j(t)$ is the measured value of the x_j variable at each time "t" and x_{j0}. is the initial condition for each x_j. Two objective functions are defined: (i) maximization of biomass productivity and (ii) minimization of the levels of substrate (in this case, extracellular sucrose). These objective functions are presented in Equations (8) and (9).

$$h_1(x_0, ESUC_0) = x_b/t_b \tag{8}$$

$$h_2(x_0, ESUC_0) = |ESUC - ESUC_0| \tag{9}$$

Where x_0 and $ESUC_0$ are the decision variables, initial conditions for inoculum and extracellular sucrose concentrations, respectively. t_b corresponds to the time in which maximum productivity is obtained. x_b corresponds to the biomass value at time t_b and ESU corresponds to the extracellular sucrose value at time t_b.

3.3 Selection of the Optimization Algorithm

The optimization problem is completely formulated in Equations (10) and (11).

$$Maximize\ h_1(x_0, ESUC_0) \tag{10}$$

$$Minimize\ h_2(x_0, ESUC_0) \tag{11}$$

where, the boundaries for x_0 and $ESUC_0$ are defined in Equations (12) and (13) which are based on expert criterion for the case of plant cell suspension cultures of *T. peruviana*,

$$2 \leq x_0 \leq 4.5 \tag{12}$$

$$13 \leq ESUC_0 \leq 28 \tag{13}$$

Finally, Equation (14) describes the dynamics of the model. For solving the optimization problem, a Multi-Objective Genetic Algorithm (MOGA) was used.

$$\frac{dX}{dt} = \sum_j r_{ij} v_j - \mu x_i \tag{14}$$

3.4 Selection of the Best solution from the Pareto front

The Pareto front found when solving the problem (stated at (10)-(14)) is shown in Figure 3, and some points

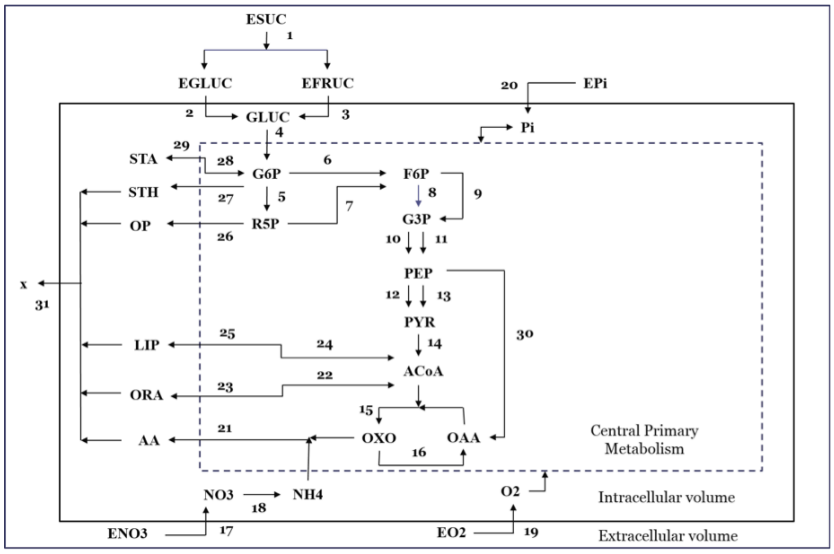

Figure 2 Model description taken from the work by [8]

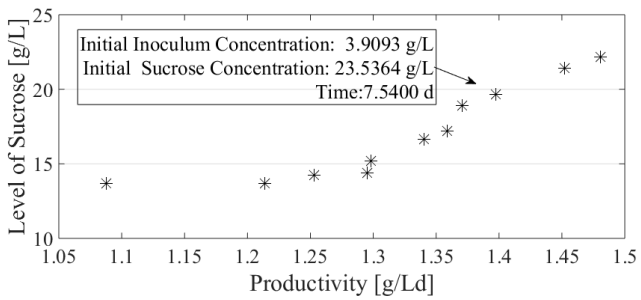

Figure 3 Pareto front for plant cell suspension cultures of *T. peruviana*

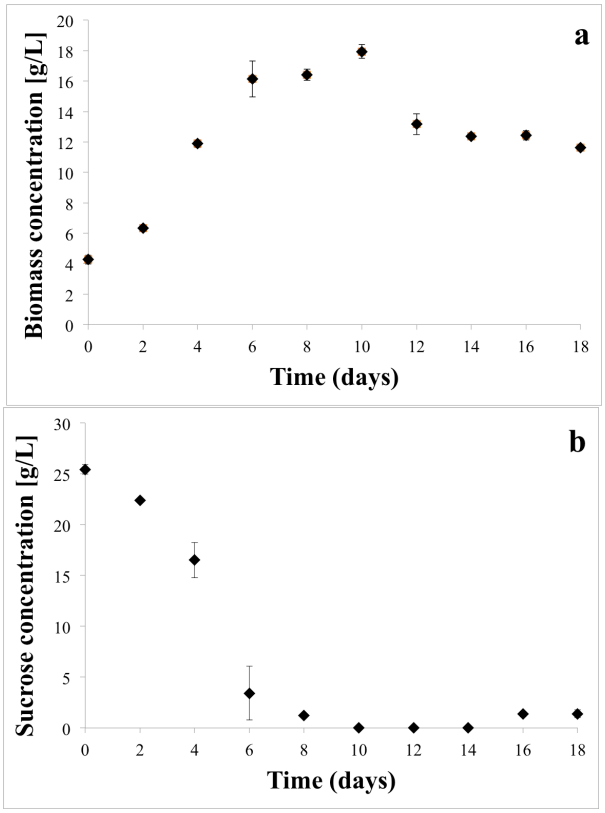

Figure 4 Validation of the selected Optima using Initial conditions of 4.27g/L of initial inoculum concentration and 25.44g/L of the initial sucrose concentration: a) Biomass Concentration and b) Sucrose concentration

are presented also in Table 1. A "Knee" solution can be observed in the region with productivities between 1.4g/Ld and 1.6g/Ld. The selected point, around to the "knee" solution, corresponds to an initial inoculum concentration of 3.9063g/L and initial sucrose concentration of 23.6289g/L to obtain a productivity of 1.5659g/Ld.

3.5 Validation of the selected solution

For validating the selected optimal point from the Pareto front (number 3 in Table 1), an experimental run was carried out by using initial concentrations of 4.27g/L of initial inoculum concentration and 25.44g/L of initial sucrose concentration (closer as possible to the selected optima due to some deviations associated to dilution effects, and impurities in the sucrose used). Results of the experimental run for biomass production in order to validate the Pareto solution are presented in Figure 4.

Experimental validation resulted in a productivity of 1.52g/Ld, reached at day 8th. This "experimental productivity" has an error of 3.1% with respect to

Table 1 Efficient optimum profile solutions in plant cell suspension cultures of *T. peruviana*

No. Initial inoculum concentration [g/L]	Initial sucrose concentration [g/L]	Productivity [g/Ld]	Time [d]	
1	3.9	13.6	1.4	4.8
2	2.0	13.6	0.7	4.8
3	**3.9**	**23.6**	**1.5**	**7.5**
4	3.9	15.8	1.4	5.3
5	3.8	22.6	1.5	7.2
6	2.0	13.6	0.7	4.8
7	3.9	17.7	1.4	5.8
8	3.9	26.4	1.7	8.4
9	3.3	13.6	1.1	4.8

the productivity value reported by the solution of the multi-objective optimization problem.

4. Conclusions

Model based multi-objective optimization is proposed as a tool for finding initial conditions (inoculum and sucrose concentrations) that optimizes the batch operation of plant cell suspension cultures of *Thevetia peruviana*. Maximization of biomass while keeping lower values of sucrose uptake are the objectives used in the multi-objective framework. Optimization results were experimentally validated, showing a god agreement. The proposed multi-objective optimization framework can be used for including maximization of metabolites production during the development of research studies in plant cell cultures of this and another species, as well as in other kind of biotechnological applications.

5. Acknowledgements

The authors thank to the Universidad de Antioquia, Universidad Nacional de Colombia, Medellin campus and LSU Agricultural Center. Adriana Villegas is grateful to the Universidad Cooperativa de Colombia for the financial support through the grant INV1879.

References

[1] O. H. Sendín, J. Vera, N. V. Torres, and J. R. Banga, "Model based optimization of biochemical systems using multiple objectives: a comparison of several solution strategies," *Mathematical and Computer Modelling of Dynamical Systems*, vol. 12, no. 5, pp. 469–487, Feb. 2007.

[2] J. O. Sendín, O. Exler, and J. R. Banga, "Multi-objective mixed integer strategy for the optimisation of biological networks," *IET Syst. Biol.*, vol. 4, no. 3, pp. 236–248, May. 2010.

[3] A. F. Villaverde, S. Bongard, K. Mauch, E. Balsa, and J. R. Banga, "Metabolic engineering with multi-objective optimization of kinetic models," *J Biotechnol.*, vol. 222, pp. 1–8, Mar. 2016.

[4] J. Vera, P. de Atauri, M. Cascante, and N. V. Torres, "Multicriteria optimization of biochemical systems by linear programming: application to production of ethanol by *Saccharomyces cerevisiae*," *Biotechnol. Bioeng.*, vol. 83, no. 3, pp. 335–343, Aug. 2003.

[5] J. Branke, K. Deb, K. Miettinen, and R. Slowinski, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, 1st ed. Berlín, Germany: Springer Berlin Heidelberg, 2008.

[6] F. Wang and J. Sheu, "Multiobjective parameter estimation problems of fermentation processes using a high ethanol tolerance yeast," *Chemical Engineering Science*, vol. 55, no. 18, pp. 3685–3695, Sep. 2000.

[7] R. Brunet, G. Guillén, and L. Jiménez, "Cleaner design of single-product biotechnological facilities through the integration of process simulation, multiobjective optimization, life cycle assessment, and principal component analysis," *Industrial & Engineering Chemistry Research*, vol. 51, no. 1, pp. 410–424, Nov. 2012.

[8] A. Villegas, J. P. Arias, D. Aragón, S. Ochoa, and M. Arias, "Structured model and parameter estimation in plant cell cultures of *Thevetia peruviana*," *Bioprocess and Biosystems Engineering*, vol. 40, no. 4, pp. 573–587, Apr. 2017.

[9] A. Villegas, J. P. Arias, D. Aragón, S. Ochoa, and M. Arias, "First principle-based models in plant suspension cell cultures: a review," *Critical Reviews in Biotechnology*, vol. 37, no. 8, pp. 1077–1089, Apr. 2017.

[10] C. A. Coello, G. B. Lamont, and D. A. van Veldhuizen, *Evolutionary Algorithms for Solving Multi-Objective Problems*, 2nd ed. New York, USA: Springer US, 2007.

[11] J. Handl, D. B. Kell, and J. Knoeles, "Multiobjective optimization in bioinformatics and computational biology," *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, vol. 4, no. 2, pp. 279–292, Apr. 2017.

[12] J. Branke, K. Deb, H. Dierolf, and M. Osswald, "Finding knees in multi-objective optimization," in *Parallel Problem Solving from Nature - PPSN VIII, 8th International Conference, Birmingham, Proceedings*, UK, 2004, pp. 18–22.

[13] L. Rachmawati and D. Srinivasan, "Multiobjective evolutionary algorithm with controllable focus on the knees of the pareto front," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 4, pp. 810–824, Aug. 2009.

[14] J. Almquist, M. Cvijovic, V. Hatzimanikatis, J. Nielsen, and M. Jirstrand, "Kinetic models in industrial biotechnology - improving cell factory performance," *Metabolic Engineering*, vol. 24, pp. 38–60, Jul. 2014.

[15] I. C. Chou and E. O. Voit, "Recent developments in parameter estimation and structure identification of biochemical and genomic systems," *Mathematical Biosciences*, vol. 219, no. 2, pp. 57–83, Jun. 2009.