A Comparative Study of Differential Evolution Algorithms for Parameter Fitting Procedures

Carlos E. Torres-Cerna
Alma Y. Alanis
CUCEI, Universidad de Guadalajara
Blvd. Marcelino Garcia Barragan,
Col. Olimpica, C.P. 44330,
Zapopan, Jalisco, Mexico.

Ignacio Poblete-Castro
Biosystems Engineering Laboratory,
Center for Bioinformatic and
Integrative Biology
Universidad Andrés Bello
Republica 239, Santiago, Chile.

Marta Bermejo-Jambrina
Esteban A. Hernandez-Vargas
Systems Medicine of Infectious Diseases
Helmholtz Centre for Infection Research
Inhoffenstr.7, 38124 Braunschweig, Germany.
Email: esteban@systemsmedicine.de

***Abstract*—Parameter fitting consists on the estimation of model parameters using experimental data from the studied process, which can be considered as a nonlinear optimization problem. In this sense, evolutionary computation has shown its great capability to solve multimodal nonlinear optimization problems. This paper compares different variants of the Differential Evolution (DE) algorithm to minimize the residual sum of squares between the outcome of the mathematical model and experimental data. To compare the different variants of the DE algorithm, a biopolymer production model is considered. Simulations results suggest a trend for the best fit using the DE/best/ variants. However, the DE/rand/ variant provides more stable results respect to the average and standard deviation of different trials. Finally, the biopolymer production problem is discussed.**

Keywords—*Parameter Estimation, Differential Evolution Algorithm, Bioprocesses*

I. Introduction

Model simulations can initiate predictions for further and conclusive experiments. This framework has the potential to advance knowledge in different engineering applications [1]–[3]. However, before considering mathematical models for reliable predictions, the values for unknown model parameters need to be estimated. Parameter procedures consist on minimizing the error between the model output and experimental data. Due to nonlinear features and complexity of several mathematical models, evolutionary optimization algorithms with features of collective intelligence are the perfect tools to address such problems [4].

A popular algorithm that has been used in several modeling works (*e.g.* [5]–[10]) to estimate model parameters is the DE algorithm [11]. For instance, in [5] parameter estimation is assessed for a nonlinear biochemical dynamic model formed by 8 differential equations; in [7] the DE algorithm is used for optimization of the feeding trajectories in fed-batch fermentation processes, and in [6] the DE algorithm is used for parameter optimization in a gene regulatory network. In these three cases, the DE algorithm showed satisfactory results.

To extend the parameter fitting discussion [5], three variants of DE algorithm are tested, these are the DE/best/1/bin [4], DE/rand/1/bin [12], and DE/best/1/exp [13]. For numerical examples, we consider the plastic production problem which is very relevant in our society. Plastics are one of the most used materials worldwide and their production rate is increasing over the years. Nowadays, worldwide plastic material consumption is approximately 140 million tons per year causing large amounts of waste which contribute to the pollution of the environment [14]. Plastic materials have been for long time produced through petrochemicals routes. Theses have a high durability and are resistant to degradation. Such properties have been for a long time considered relevant for the industrial production. However, plastics are a large source of environmental and waste management problems nowadays [14].

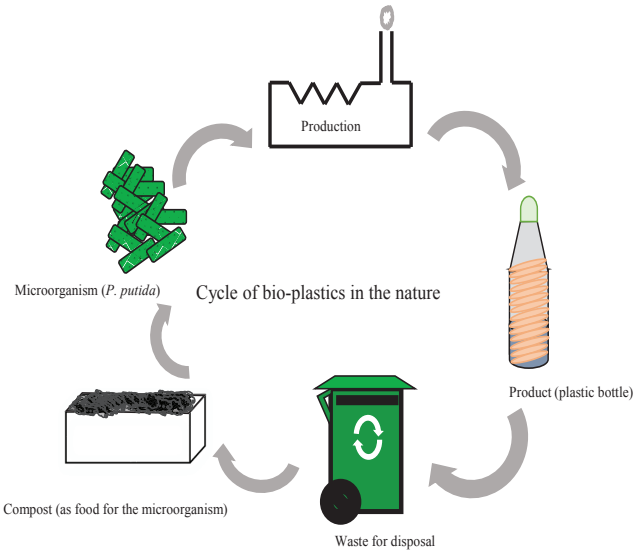

Fig. 1. **Bio-plastics life cycle.** Bio-plastics are produced by microorganisms, which are collected after being used as solid wasters. Consequently, these are degraded and used as carbon sources for the microorganisms.

The replacement of a fraction of synthetic plastics with biodegradable polymers produced from renewable resources offers to impact on the overall consumption of fossil fuels, environmental pollution and solid waste management [15]. Bio-plastics can easily be degraded (Fig.1), in this way the

plastics abandoned in the nature could be largely degraded, serving as a carbon substrate to different microorganisms.

In this work, a mathematical model that represents the bio-plastic production will be used to test the performance of DE algorithm variants. This paper is organized as follow: in section II, the mathematical model for bio-polymer production is introduced. In section III the objective functions used to evaluate the estimation results are explained. In section IV the variants of the DE algorithm are presented. Results are discussed in section V, and conclusions are presented in section VI.

II. Mathematical model

Among the candidates for bio-plastics production, Poly-hydroxyalkanoates (PHAs) have been drawing much attention because of their similar material properties to conventional plastics and complete biodegradability. In addition, PHAs are naturally produced as granules in the cytoplasm of cells. Nevertheless, the amount of bacterial production is until today insufficient compared to the demand and the reduced quantity of production, influencing the price of bio-plastics in the market. Mathematical modeling represents an important venue to reduce the costs of bio-polymer production.

To this end, the mathematical model proposed here is described by equations (1)-(4), where S is the glucose as carbon source, N is the NH_4 concentration, X is the active biomass, and P is the biopolymer PHA.

$$\frac{dS(t)}{dt} = -k_1\mu_S(t) \quad (1)$$

$$\frac{dN(t)}{dt} = -k_2\mu_N(t) \quad (2)$$

$$\frac{dX(t)}{dt} = \mu(t) \quad (3)$$

$$\frac{dP(t)}{dt} = k_3\mu(t) + k_4\mu_P(t) \quad (4)$$

In a similar way as [16], the limiting nutrient is NH_4 and the total biomass (CDW) is assumed to be composed by the catalytically active biomass and the product PHA, thus $CDW = X + P$. The substrate S is consumed mainly by the active biomass with a rate constant k_1, which is modeled by equation (1) and the following Monod function:

$$\mu_S(t) = \frac{X(t)}{X(t) + K_S}S(t) \quad (5)$$

where K_S is the saturation or affinity constant of the substrate S. NH_4 dynamics are represented by the equation (2). Experimental data in [17] indicate that the ammonium is completely consumed in the growth phase. Note that we tested different functions for the NH_4 consumption, nevertheless, the best estimations were achieved by the following rate equation:

$$\mu_N(t) = \frac{X(t)}{X(t) + K_N}N(t) \quad (6)$$

where K_N is the saturation or affinity constant of the substrate N. Active biomass dynamics by the *Pseudomonas putida* are presented in equation (3). The growth rate of biomass with glucose is defined according to a "double Monod" relation. The limiting substrate ammonium (N) is essential to produce the active biomass X and limits its synthesis at low concentrations. The active biomass growth μ writes as follows:

$$\mu(t) = \mu_{max}\left(\frac{S(t)}{S(t) + K_{S1}}\right)\left(\frac{N(t)}{N(t) + K_{N1}}\right)X(t) \quad (7)$$

where μ_{max} is the maximum growth rate, this is set by the experimental values presented in [17]. The bioplymer PHA is produced with a rate k_4 by the active biomass dynamics (X), which is trigger by a limited amount of ammonium provided in the growth medium, this can be expressed by the following equation:

$$\mu_p(t) = \frac{K_P}{N(t) + K_P}X(t) \quad (8)$$

Note that $\mu_p(t)$ term will increase when the value of ammonium decreases. To provide a better fitting for the PHA, we needed to introduce a growing term based on the biomass growth k_3 implying that small amounts of PHA are produced during biomass growth. Thus, the model parameter set to be estimated by the DE variants reads as follows:

$$\theta = [k_1, k_2, k_3, k_4, K_S, K_N, K_{S1}, K_{N1}, K_P] \quad (9)$$

III. Cost function

Parametric estimation for mathematical models can be understood as the values search for the model parameters (θ) that minimizes the difference between the outcome of model $\bar{y}_i$ and experimental data y_i (inverse problem). This search is restricted by the system dynamics and process constraints. Due to the different scales of experimental values, we considered the sum of the square of weighted residues *SSWR*, this writes as follows:

$$SSWR = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(\frac{y_i^j - \bar{y}_i^j}{max(y_i^j)}\right)^2 \quad (10)$$

where m is the total number of variables, n is the total number of experimental data for each variable j, y_i is the *i-th* experimental data point from [17] and $\bar{y}_i$ is the *i-th* outcome of the mathematical model described by equations (1)-(4).

IV. Differential Evolution algorithm

The minimization of the cost function (10) implies a non-linear optimization problem with several variables, which can be tackled using the DE algorithm [11]. The DE algorithm is a population-based optimization algorithm, where each individual in the population is a n-dimensional vector that represents a candidate solution to the problem. The individuals can be defined as:

$$x_{i,g} = [\chi_{i,g}^1, \chi_{i,g}^2, ..., \chi_{i,g}^n] \quad (11)$$

with

$$\begin{aligned} &i = \{1, 2, ..., NP\}\text{, and} \\ &g = \{1, 2, ..., G\} \end{aligned} \quad (12)$$

where $\chi_{i,g}^{n}$ are the elements of the *i*-th individual in the generation g, n is the dimension of the problem, the variable G represents the maximum number of generations and NP is the size of the population. The basic idea behind the DE algorithm is that two individuals mutually excluding $x_1 \neq x_2$ are picked randomly among population, its difference is scaled and added to a third individual $x_3 \notin \{x_1, x_2\}$ chosen randomly to create a new mutant vector v_i. The DE algorithm is explained with the following steps:

I) Initialization: the population is initialized within a bounded search space defined as $\chi_{min} = \{\chi_{min}^1, ..., \chi_{min}^n\}$ and $\chi_{max} = \{\chi_{max}^1, ..., \chi_{max}^n\}$, where X_{min} and X_{max} are the sets of lower and upper bounds for the elements of the individuals.

II) Mutation: different strategies exist for the mutation procedure. In this work, the following mutation strategies are considered [11], [18]:

1) DE/best/1/

$$v_{i,g} = x_{best,g} + F(x_{r1,g} - x_{r2,g}) \tag{13}$$

2) DE/rand/1/

$$v_{i,g} = x_{r3,g} + F(x_{r1,g} - x_{r2,g}) \tag{14}$$

where $v_{i,g}$ is the mutant vector; $x_{best,g}$ is the individual with the best fit in the current generation; $x_{r1,g}, x_{r2,g}$ and $x_{r3,g}$ are different vectors chosen randomly among the population, and $F \in [0, 1]$ is the scaling factor.

III) Crossover: a new trial vector is generated recombining the mutant $v_{i,g}$ and target $x_{i,g}$ vectors. The principal approaches of crossover are the binomial and exponential [11], [18], described in equations (15) and (16) respectively.

$$u_{i,g}^{j} = \begin{cases} v_{i,g}^{j} & if \quad r_{cj} \leq Cr \ or \ j = j_{rand} \\ x_{i,g}^{j} & otherwise \end{cases} \tag{15}$$

where $j \in [1, n]$ is the j-th component of $u_{i,g}$, $v_{i,g}$, and $x_{i,g}$, $r_{cj} \in [0, 1]$ is a random number taken from a uniform distribution. $Cr \in [0, 1]$ is a constant crossover rate, and $j_{rand} \in [1, n]$ is a random number taken from a uniform distribution.

$$u_{i,g}^{j} = \begin{cases} v_{i,g}^{j} & if \quad j \in A \cup B \\ x_{i,g}^{j} & otherwise \end{cases} \tag{16}$$

where $A := \{s, ..., min(n, s + L - 1)\}$, and $B := \{1, ..., s + L - n - 1\}$. s represents the starting element in a target vector to create a new trial vector recombining the L elements donated by the mutant vector. s is an integer randomly taken from the interval $[1, n]$, and L is draw from the interval $[1, n]$ according with the following code:

```
L = 0
while ((rand(0,1) ≤ Cr) and (L < n)) do
    L = L + 1
end while
```

where Cr is called the crossover rate. F, NP and Cr are the control parameters of the DE algorithm.

VI) Selection: the cost function is used to determinate which vector survives for the new generation $(g + 1)$, the mutant $v_{i,g}$ and target $x_{i,g}$ vectors are evaluated in order to find which vector has the best yield, the basic idea is presented next:

$$x_{i,g+1} = \begin{cases} u_{i,g} & if \quad f(u_{i,g}) \leq f(x_{i,g}) \\ x_{i,g} & if \quad f(x_{i,g}) < f(u_{i,g}) \end{cases} \tag{17}$$

where $f(\bullet)$ is the evaluation of the objective function.

Note that some characteristics from a generation can be transferred to the next one in the DE/best/ variants. In this way, the members of the next generation can have certain properties of the current leader, reducing the space of search. Otherwise, the next generation is randomly generated favoring the space of search (DE/rand/).

V. Results

The parameters of the mathematical model equations (1)-(4) can not be measured directly using common model optimization strategies (*e.g.* least value, steepest descendant, inner and outer approximation) due to the non-linearity and complexity of the model. The DE algorithm is an algorithm with self-adaption, self-organizing and self-learning which allows to address challenging problems that can not be solved by traditional methods.

The three control parameters are setting as, $Cr = 0.7$, $F = 0.8$, and population size depends on the number of parameters to be estimated $NP = 20 * n$.

A. DE algorithm performance

The DE variants were assessed to estimate the model parameters presented in (9) with the experimental data of PHA production by the wild-type strain *P. putida* [17]. Table I presents the performance of the three variants after 500 trials.

TABLE I
DE algorithm variants. SD represents the standard deviation.

Variant	Best	Average ± SD	Time (sec)
DE/Best/1/exp	0.0711	0.2049 ± 0.1308	865
DE/Best/1/bin	0.0709	0.2560 ± 0.1693	927
DE/Rand/1/bin	0.0847	0.1099 ± 0.0448	930

In general, the three variants reveal a good performance, and its outcome provides a close approximation to experimental data. We can observe in Table I that DE/best/ variants perform better in the "best fitting" columns respect to DE/Rand/. This can be attributed to the fast convergence time, see Fig. 2. However, it is interesting to observe in this example that the intuitive idea of transferring to the next generation certain properties of the current leader by the variants DE/best/ does not contribute to the average reduction of the error but increasing the standard deviation. Note that for several reasons, experimental data possess large variations, therefore robust algorithms to estimate parameters are highly preferred.

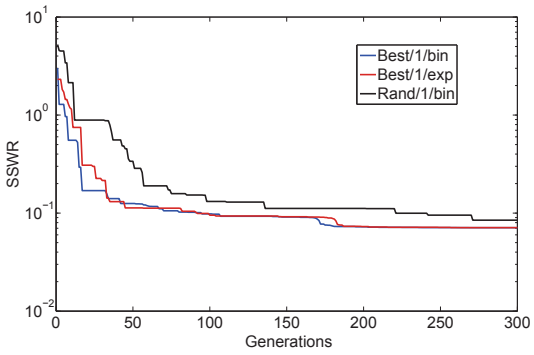

Fig. 2. **Estimation error dynamics.**

B. Biological problem

The experimental work in [17] assessed the growth and PHA production performance by the wild type *P. putida* KT2440 in minimal medium using glucose as a carbon source (18.5 g L^{-1}) and a limiting amount of ammonium as nitrogen source 350 (mg L^{-1}). The experimental results by [17] are shown in Fig. 3. Cells (X) present an exponential growth during the first 20 hours, subsequently the cell growth slows down which is mainly attributed to the limited levels of glucose and ammonium presented in the medium. The PHA content exhibits two phases; the first is proportional to the cell growth and the second is when ammonium presented in the medium is less than 0.1 (g L^{-1}).

Here, the reactor dynamics represented with the system (1)-(4) is portrayed in Fig. 3, showing the good fitting to the experimental data (empty circle) presented in [17]. Note that previous mathematical models for biopolymer production proposed more complex equations [19]–[22]. Here, different terms were considered for the substrates equations (1)-(2), *e.g.* bilinear terms (k_1SX) or more complex like Hill functions ($X^n/(X^n + K_S)$), however further improvements in the fitting values were not achieved. Thus, these results show the importance of a minimalist model.

To achieve good predictions for PHA dynamics, a production term as a function of the bacterial growth (k_3) was needed. The highest production to PHA can be also attributed to the adequate consumption of the substrates. Fig. 3 revealed that the growth of the biomass depends directly on the ammonium. We can see when the ammonium is exhausted in medium, the biomass stops growing. Furthermore, both substrate and ammonium decrease logarithmically. The substrate and ammonium dynamics are simpler than the biomass and PHA production, due to they have a linear behavior in a part of the process. On the other hand, biomass and PHA dynamics are more complex throughout the process, resulting only in more complicated dynamics to model.

The principal advantage of our model in comparison with previous mathematical models is that our model is less complex. For example, in [19]–[24] proposed substrates equations with bilinear terms (k_1SX) or more complex like Hill functions ($X^n/(X^n + K_S)$), however further improvements in the cost function were not achieved.

VI. Conclusions

In this paper, we presented a comparative study of different variants of DE algorithm tested for parametric estimation of a mathematical model. This estimation presents a good fitting for the *P. putida* strain. The DE algorithm showed to be a very useful tool for parameter estimation, something that we could not reach by applying traditional methods. Comparing the different variants of the DE algorithm, simulations results revealed a trend for the DE/best/, which also offers more restricted search space for parameter values. However, for this example, DE/rand/ performs better respect to the average and standard deviation.

Further identifiability and experimental studies need to be developed in order to corroborate the parameter values. Future work will compare to other optimization methods, like Particle Swarm Optimization (PSO), or Simulated Annealing (SA). In addition, the extension of the present model to the fed-batch case will be relevant to propose control strategies to minimize production costs.

Acknowledgment

This work was supported by CONACyT Mexico and DAAD Germany through the program PROALMEX funding the project "OPTREAT"; and iMed - the Helmholtz Initiative on Personalized Medicine.

References

[1] A. Quarteroni, "Mathematical models in science and engineering," *Notices of the AMS*, vol. 56, no. 1, pp. 10–19, 2009.

[2] E. A. Hernandez-Vargas and R. H. Middleton, "Modeling the three stages in HIV infection," *Journal of theoretical biology*, vol. 320, pp. 33–40, 2013.

[3] A. Boianelli, E. Pettini, G. Prota, D. Medaglini, and A. Vicino, "A Stochastic Model for CD4+ T Cell Proliferation and Dissemination Network in Primary Immune Response," *PloS one*, vol. 10, no. 8, p. e0135787, jan 2015.

[4] D. Simon, *Evolutionary optimization algorithms*. John Wiley & Sons, 2013.

[5] C. G. Moles, P. Mendes, and J. R. Banga, "Parameter estimation in biochemical pathways: a comparison of global optimization methods," *Genome research*, vol. 13, no. 11, pp. 2467–2474, 2003.

[6] N. Noman and H. Iba, "Inference of gene regulatory networks using s-system and differential evolution," in *Proceedings of the 7th annual conference on Genetic and evolutionary computation*. ACM, 2005, pp. 439–446.

[7] M. Rocha, R. Mendes, O. Rocha, I. Rocha, and E. C. Ferreira, "Optimization of fed-batch fermentation processes with bio-inspired algorithms," *Expert Systems with Applications*, vol. 41, no. 5, pp. 2186–2195, 2014.

[8] E. A. Hernandez-Vargas, E. Wilk, L. Canini, F. R. Toapanta, S. C. Binder, A. Uvarovskii, T. M. Ross, C. a. Guzmán, A. S. Perelson, and M. Meyer-Hermann, "Effects of aging on influenza virus infection dynamics," *Journal of Virology*, vol. 88, no. 8, pp. 4123–31, apr 2014.

[9] V. K. Nguyen and E. A. Hernandez-vargas, "Identifiability Challenges in Mathematical Models of Viral Infectious Diseases," *IFAC-PapersOnLine*, vol. 48, no. 28, pp. 2–7, 2015.

[10] A. Boianelli, V. K. Nguyen, T. Ebensen, K. Schulze, E. Wilk, N. Sharma, S. Stegemann-Koniszewski, D. Bruder, F. R. Toapanta, C. a. Guzmán, M. Meyer-Hermann, and E. A. Hernandez-Vargas, "Modeling Influenza Virus Infection: A Roadmap for Influenza Research," vol. 7, no. 10, pp. 5274–5304, 2015.

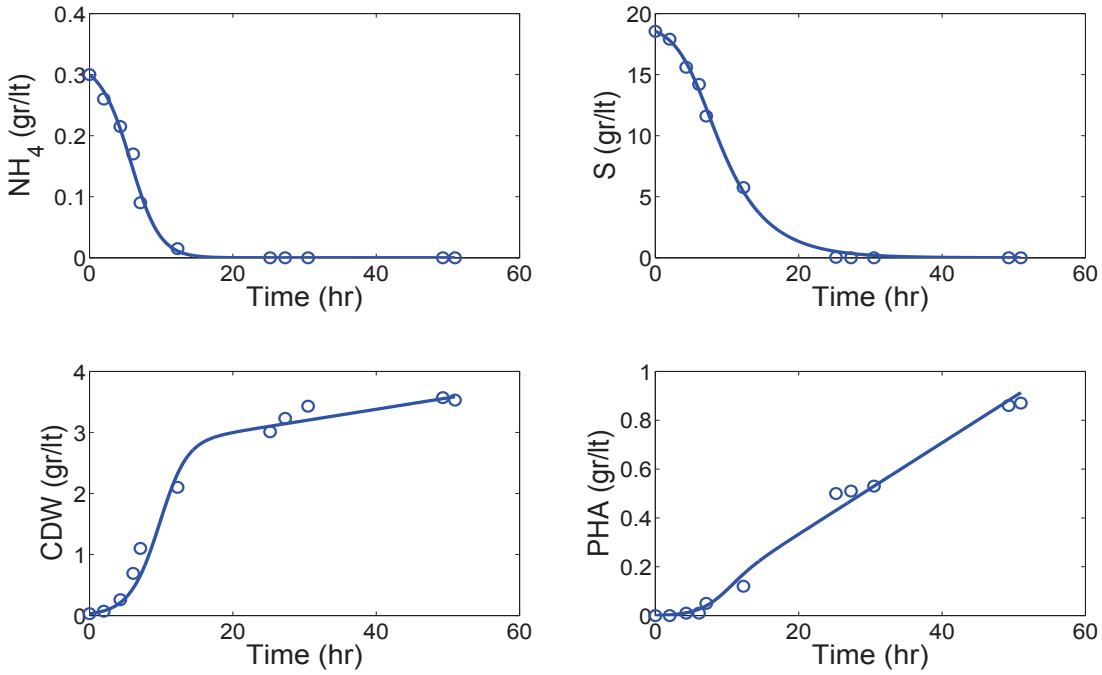

Fig. 3. **Simulation results for the main variables during biopolymer production.** S is the glucose as carbon source, N is the NH_4 concentration, *CDW* is the total biomass and P is the biopolymer PHA. Empty circles represent the corresponding experimental data from [17].

[11] R. Storn and K. Price, *Differential evolution-a simple and efficient adaptive scheme for global optimization over continuous spaces*. ICSI Berkeley, 1995, vol. 3.

[12] J. Montgomery and S. Chen, "An analysis of the operation of differential evolution at high and low crossover rates," in *Evolutionary Computation (CEC), 2010 IEEE Congress on*, July 2010, pp. 1–8.

[13] S. Das and P. N. Suganthan, "Differential evolution: a survey of the state-of-the-art," *Evolutionary Computation, IEEE Transactions on*, vol. 15, no. 1, pp. 4–31, 2011.

[14] R. Zhao, P. Torley, and P. Halley, "Emerging biodegradable materials: starch- and protein-based bio-nanocomposites," *Journal of Materials Science*, vol. 43, no. 9, pp. 3058–3071, 2008.

[15] S. Lee, "Bacterial polyhydroxyalkanoates," *Biotechnology and bioengineering*, vol. 49, pp. 1–14, 1996.

[16] S. Khanna and A. K. Srivastava, "Optimization of nutrient feed concentration and addition time for production of poly(β-hydroxybutyrate)," *Enzyme and Microbial Technology*, vol. 39, no. 5, pp. 1145–1151, Sep. 2006.

[17] I. Poblete-Castro, D. Binger, A. Rodrigues, J. Becker, V. a. P. Martins Dos Santos, and C. Wittmann, "In-silico-driven metabolic engineering of Pseudomonas putida for enhanced production of poly-hydroxyalkanoates," *Metabolic engineering*, vol. 15, pp. 113–23, Jan. 2013.

[18] R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," *Journal of global optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[19] P. Raje and A. Srivastava, "Updated mathematical model and fed-batch strategies for poly-β-Hydroxybutyrate (PHB) production by Alcaligenes eutrophus," *Bioresource technology*, vol. 8524, no. 97, pp. 185–192, 1998.

[20] K. Johnson, R. Kleerebezem, and M. C. M. van Loosdrecht, "Model-based data evaluation of polyhydroxybutyrate producing mixed microbial cultures in aerobic sequencing batch and fed-batch reactors." *Biotechnology and bioengineering*, vol. 104, no. 1, pp. 50–67, Sep. 2009.

[21] P. Patwardhan and a.K Srivastava, "Model-based fed-batch cultivation of R. eutropha for enhanced biopolymer production," *Biochemical Engineering Journal*, vol. 20, no. 1, pp. 21–28, Aug. 2004.

[22] P. Horvat, I. Vrana Špoljarić, M. Lopar, A. Atlić, M. Koller, and G. Braunegg, "Mathematical modelling and process optimization of a continuous 5-stage bioreactor cascade for production of poly[-(R)-3-hydroxybutyrate] by Cupriavidus necator." *Bioprocess and biosystems engineering*, Nov. 2012.

[23] S. Khanna and A. Srivastava, "A simple structured mathematical model for biopolymer (PHB) production," *Biotechnology progress*, no. 15, pp. 830–838, 2005.

[24] S. Khanna and A. K. Srivastava, "Optimization of nutrient feed concentration and addition time for production of poly(β-hydroxybutyrate)," *Enzyme and Microbial Technology*, vol. 39, no. 5, pp. 1145–1151, Sep. 2006.