Direct Sequential Based Firefly Algorithm for the α-Pinene Isomerization Problem

Ana Maria A.C. Rocha[1], Marisa C. Martins[1], M. Fernanda P. Costa[2], and Edite M.G.P. Fernandes[1]

[1] Algoritmi Research Centre,
University of Minho, 4710-057 Braga, Portugal
arocha@dps.uminho.pt, marisacunhamartins@gmail.com, emgpf@dps.uminho.pt
[2] Centre of Mathematics, University of Minho, 4800-058 Guimarães, Portugal,
mfc@math.uminho.pt

Abstract. The problem herein addressed is a parameter estimation problem of the α-pinene process. The state variables of this bioengineering process satisfy a set of differential equations and depend on a set of unknown parameters. A dynamic system based parameter estimation problem aiming to estimate the model parameter values in a way that the predicted state variables best fit the experimentally observed state values is used. A numerical direct method, known as direct sequential procedure, is implemented giving rise to a finite bound constrained nonlinear optimization problem, which is solved by the metaheuristic firefly algorithm (FA). A Matlab^{TM} programming environment is developed with the mathematical model and the computational application of the method. The results produced by FA, when compared to those of the `fmincon` function and other metaheuristics, are competitive.

Keywords: α-pinene isomerization, parameter estimation, direct sequential procedure, Firefly algorithm

1 Introduction

Bioinformatics expertise is usually required to address some issues related with slow and expensive processes from the biotechnology area. This type of skills may be used to understand the behavior of certain organisms or biochemicals, such as the case of the parameter estimation problem on the α-pinene isomerization dynamic model. A parameter estimation problem aims to find the parameter values of a mathematical model that gives the best possible fit with existing experimental data [1]. Thus, in a dynamic model based parameter estimation process, an objective functional that gives the mean squared error between the model predicted state values and the observed values (within a fixed time interval) is minimized, subject to a set of differential equations. In a dynamic system, the parameter estimation concept may lead to confusion since the parameters are not usually time-dependent in the mathematical model. Nevertheless, in a problem like this, the parameters are the decision variables of the optimization process, although they are constant throughout the entire simulation process.

The general form of a first order ordinary differential equation (ODE) is given by

$$\frac{dy}{dt} = g(t, y) \tag{1}$$

where t is the independent variable (here, it is assumed that t represents time) and $y \equiv y(t)$ is the dependent variable. The equation is called ODE since the unknown function depends on a single independent variable and the order of the ODE is defined by the highest derivative order that appears in the equation. The equation is called linear if the function $g(t, y)$ is linear on y. Further, a solution of the ODE is a function $y(t)$ that satisfies the equation (1) for all values of t in the domain of y. Most of the time, we are not interested in all solutions to a differential equation, but only in a particular solution satisfying an extra condition. If the extra condition provides the value of y at the initial instant of time, the problem is called 'initial value problem'. Hereinafter the more compact notation $y'(t) = dy/dt$ is used.

Numerical methods for solving ordinary differential equations are discretization methods that compute approximations to the solution $y(t)$ at a finite set of points $t_0, t_1, t_2, \ldots, t_n$ of the independent variable interval $t_i \leq t \leq t_f$, where t_i and t_f are the initial and final values of the interval, respectively. A variety of numerical methods for solving a system of ODEs is available in the literature, being the most known the Euler's method, the Runge-Kutta method and the backward differentiation formula [2].

A dynamic mathematical model emerges when the optimization and control of industrial bioprocesses [3] are carried out. The construction of the model involves several stages. First, the objectives are defined based on theoretical and/or empirical knowledge of the process under study. In general, the mathematical model depends on a set of unknown parameters that require to be investigated. Second, the values for the parameters are estimated based on experimental data, assuming that the mathematical model simulates the process the best possible way. The goal of the parameter estimation problem is to calibrate the model so that it can reproduce the experimental results as close as possible. This is performed by minimizing an objective function that measures the goodness of the fit. Finally, the model with the estimated parameter values may be validated [4,5].

In general, the mathematical modeling of bioprocess engineering problems involves nonlinear dynamic equations as constraints and a nonlinear objective function to be optimized, giving the so called dynamic optimization (DO) problem. Nonconvex and multimodal functions frequently arise in DO. Moreover, some kind of noise and/or discontinuities may be present making the problem even more complex. Therefore, robustness and efficiency are crucial properties of an optimization solver so that it is capable of computing a good approximation to a global solution to the DO problem without an excessive computational effort.

Gradient-based local methods are generally very efficient to solve finite-dimensional constrained nonlinear programming (NLP) problems, although they

can only deal with problems where the objective function and the constraint functions are continuously differentiable [6]. In general, they may lead to suboptimal solutions if multiple local optima are present.

The alternative to compute a global optimal solution to a NLP problem is to use a global optimization method [1,7,8,9,10]. There has been a growing interest in developing algorithms that converge to the global optimal solution. Global search algorithms often emerge from the modifications introduced in local search algorithms. A common practice is to introduce some stochastic components to force the algorithm to diversify the search for other solutions and explore the search space. Then, they become stochastic algorithms. The stochastic component can be introduced in various form, such as for example, a simple random generation of solutions in the search space, and the use of random numbers to define the path of solutions. Most stochastic algorithms are also known as metaheuristics since they combine random components with historical knowledge of previous search in regions that have been explored and are in the neighborhood of local solutions. This is an important issue since premature convergence to local optimal solutions can be avoided. Metaheuristics are approximate methods or heuristics that are designed to search for good solutions, known as near-optimal solutions, with less computational effort and time than the more classical algorithms. Algorithms that simulate the behavior of animals in nature have been widely used to solve complex mathematical problems. Classical examples of this type of algorithms are the particle swarm optimization [11] and more recently the firefly algorithm (FA) [12].

Deterministic global optimization techniques have also been used to find a global solution that best reconciles the model parameters and measurements [13,14,15]. Although the convergence to the global optimal solution is guaranteed, a deterministic method is only feasible for a small number of decision variables. Since stochastic methods are able to converge rapidly to the vicinity of the global solution, although they are very expensive to locate a global solution with high quality, hybrid methods that combine global and local strategies have recently become very popular [16].

Since the problem of estimating the parameters of dynamic models of complex biological systems is becoming increasingly important, the contribution of the present study is focused on the implementation of the metaheuristic FA to solve the finite NLP problem that arises from using a numerical direct approach to locate a global optimal solution of a specific DO problem - the parameter estimation problem on the α-pinene isomerization. A powerful direct method that transcribes the dynamic model parameter estimation problem in a small finite-dimensional optimization problem through the discretization of the decision variables only is used.

The remaining part of this paper is organized as follows. In Section 2, we describe the α-pinene isomerization problem and in Section 3 we present the selected methodology for solving the α-pinene problem. Section 4 presents the results of the numerical experiments and comparisons and Section 5 contains the conclusions of the present study.

2 Isomerization of α-pinene

The α-pinene isomerization was studied in 1973 by Box, Hunter, MacGregor and Erjavec in [17] and goals the estimation of a set of five velocity parameters $p_1, \ldots, p_5$ of a homogeneous chemical reaction. The reaction scheme for this chemical reaction that describes the thermal isomerization of α-pinene, y_1, to dipentene, y_2, and allo-ocimen, y_3, which in turn produces α- and β-pyronene, y_4, and a dimer, y_5, is shown in Fig. 1.

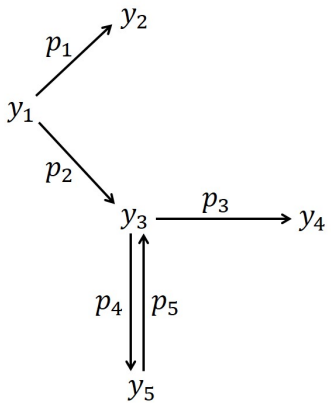

Fig. 1. α-pinene isomerization mechanism (adapted from [18]).

Fuguitt and Hawkins, in [19], observed the concentrations of the reactant and the four products in eight time instants of the interval $[0, 36420]$. After the chemical reaction orders having been defined, the formulation of the mathematical model describing the concentration of each compound over time may be set. In [20], a first order kinetics is assumed and the following system of first order ODEs to describe the dynamics of the concentration of the compounds over time is proposed:

$$\begin{cases} y_1' = -(p_1 + p_2)y_1 \\ y_2' = p_1 y_1 \\ y_3' = p_2 y_1 - (p_3 + p_4)y_3 + p_5 y_5 \\ y_4' = p_3 y_3 \\ y_5' = p_4 y_3 - p_5 y_5 \end{cases} \tag{2}$$

for $t \in [0, 36420]$, with the initial conditions

$$y_1(0) = 100,\ y_2(0) = 0,\ y_3(0) = 0,\ y_4(0) = 0,\ y_5(0) = 0. \tag{3}$$

The velocity parameters $p_1, \ldots, p_5$ are unknown but can be estimated through the minimization of an objective functional which measures the distance between the experimental observed values and the model predicted values over the entire

time interval. Thus, the formulation of α-pinene isomerization problem is the following:

$$\begin{aligned} \min_{p} \; & J(p) \equiv \int_0^{t_f} (y(t) - y^{\text{obs}}(t))^T (y(t) - y^{\text{obs}}(t)) dt \\ \text{subject to } & y'(t) = g(t, y(t), p) \quad \text{ODEs listed in (2)} \\ & p^{\text{L}} \leq p \leq p^{\text{U}} \\ & y(0) = (100, 0, 0, 0, 0) \end{aligned} \tag{4}$$

where $y(t)$ denotes the vector $(y_1(t), y_2(t), y_3(t), y_4(t), y_5(t))$ of the state variable functions (predicted by the model), J is the objective functional to be minimized, p is the vector that contains the parameters to be estimated, also referred to as the vector of the decision variables, y^{obs} contains the experimentally observed values of the state variables, p^{L} and p^{U} represent the vectors with the lower and upper bounds on the parameters, respectively, and $y(0)$ is the vector with the initial conditions. This type of parameter estimation problem belongs to a class of DO problems where the decision variables are not time-dependent.

3 Methodology

The α-pinene isomerization problem (4) consists of estimating a set of five parameters $(p_1, \ldots, , p_5)$ in a way that the state variables y_1, y_2, y_3, y_4, y_5 that satisfy the system of ODEs shown in (2), with initial conditions (3), best fit the experimentally observed state values. This dynamic system based parameter estimation problem is solved by a numerical direct method and the resulting finite bound constrained nonlinear programming (BCNLP) problem is solved to global optimality by a metaheuristic, known as FA. In the remaining part of this section, the numerical direct approach is briefly described and the main ideas behind the FA are presented.

3.1 A Direct Sequential Approach

To solve the dynamic system based parameter estimation problem, a numerical direct method is used. This type of method discretizes the problem (4) and applies nonlinear programming techniques to the resulting finite-dimensional optimization problem. Methods for solving DO problems like (4) can be classified into indirect methods and direct methods [21]. Indirect methods use the first order necessary conditions from the Pontryagin's Minimum Principle to reformulate the problem as a two-point boundary value problem. Although this is a widely used technique, the resulting boundary value problems may become difficult to solve specially if the problem contains state variable constraints [22,23]. In direct methods, the optimization (4) is performed directly. Depending on whether the system of ODEs are integrated explicitly or implicitly, two different approaches emerge, the direct sequential approach and the direct simultaneous approach, respectively. In the direct sequential approach, also denoted by single-shooting or control vector parametrization (specially if the decision variables

are time-dependent), the optimization is carried out in the space of the input variables only. Given a set of values for the decision variables, the system of ODEs are accurately integrated (over the entire time interval) using specific numerical integration formulae so that the objective functional can be evaluated. Thus, the differential equations are satisfied at each iteration of the optimization procedure [22,23,24]. This is the main characteristic of the sequential approach, also coined as 'feasible path' approach. The strategy may lead to a slow convergence process since the differential equations are solved again and again. On the other hand, the direct simultaneous approach (orthogonal collocation, complete parametrization or full discretization) is computationally less expensive since an approximation to the solution of the ODEs is used instead. The optimization itself is carried out in the full space of discretized inputs and state variables leading to a large-dimensional finite NLP problem, in particular if a fine grid of points is required to obtain high level of integration accuracy. Furthermore, the number of time stages and collocation points as well as the position of the collocation points have to be chosen before using the NLP solver. Thus, efficient NLP solvers are crucial when solving the finite optimization problem. The ODEs are satisfied only at a finite number of time instants in the interval (at the solution of the optimization problem), which is why it is called 'infeasible path' approach [21,22].

To summarize, the advantages of the direct sequential approach over the direct simultaneous are highlighted as follows:

- the NLP problem is relatively small-dimensional;
- an error control mechanism within the numerical integration formula is able to enforce the accuracy of the state variable values;
- the ODEs are satisfied at each iteration of the NLP algorithm, although this may lead to a computationally demanding process.

It is recognized that a direct sequential framework is easily constructed and may link reliable and efficient ODEs numerical integration and NLP solvers. However, its implementation requires repeated numerical integration of the ODEs. In the context of solving the herein studied optimization problem, the ODEs (model (2)) with the initial conditions (3) are solved by the Matlab$^{\text{TM}}$[3] function `ode15s`. This function implements a variable order method for solving stiff differential equations and is appropriate since the equations include some terms that might lead to dramatic changes and oscillations in the solutions, in a small time interval, when compared to the interval of integration. The direct sequential optimization process can be described schematically as shown in Fig. 2. Therefore, the direct sequential approach transcribes the problem (4) into a small finite-dimensional BCNLP problem through the discretization of the decision variables $p_1, \ldots, p_5$ only, while the system of ordinary differential equations is embedded in the NLP problem.

The experimental data provided by [19] in the interval $t \in [0, t_f]$ consider the set of eight time instants {1230, 3060, 4920, 7800, 10680, 15030, 22620,

[3] Matlab is a registered trademark of the MathWorks, Inc.

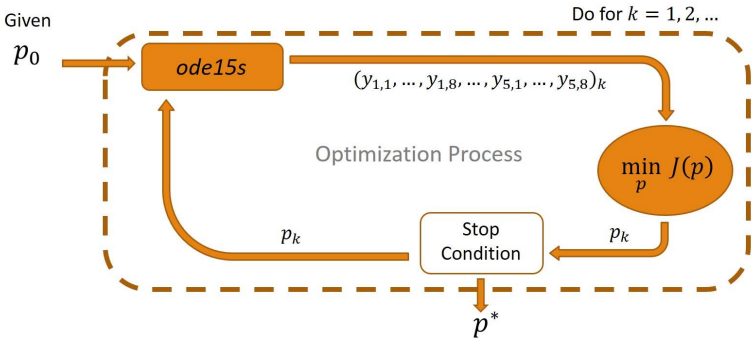

Fig. 2. Direct sequential process.

36420} and the corresponding observed concentrations of the reactant and the four products $(y_{1,i}^{\text{obs}}, y_{2,i}^{\text{obs}}, y_{3,i}^{\text{obs}}, y_{4,i}^{\text{obs}}, y_{5,i}^{\text{obs}}, i = 1, \ldots, 8)$ are shown in Table 1. Thus, based on the information herein presented, the α-pinene isomerization parameter estimation problem is formulated as

$$\begin{aligned} \min_p \; & J(p) \equiv \sum_{j=1}^{5} \sum_{i=1}^{8} \left(y_j(t_i) - y_{j,i}^{\text{obs}} \right)^2 \\ \text{subject to } & y'(t) = g(t, y(t), p) \qquad \text{ODEs listed in (2)} \\ & p^{\text{L}} \leq p \leq p^{\text{U}} \\ & y(0) = (100, 0, 0, 0, 0)\,. \end{aligned} \tag{5}$$

Table 1. Experimental values of the concentration of the reactant and the four products, for eight time instants.

	1230.0	3060.0	4920.0	7800.0	10680.0	15030.0	22620.0	36420.0
y_1^{obs}	88.35	76.4	65.1	50.4	37.5	25.9	14.0	4.5
y_2^{obs}	7.3	15.6	23.1	32.9	42.7	49.1	57.4	63.1
y_3^{obs}	2.3	4.5	5.3	6.0	6.0	5.9	5.1	3.8
y_4^{obs}	0.4	0.7	1.1	1.5	1.9	2.2	2.6	2.9
y_5^{obs}	1.75	2.8	5.8	9.3	12.0	17.0	21.0	25.7

3.2 The Firefly Algorithm

In order to avoid the convergence to a local solution of a BCNLP problem without providing a good initial approximation, a global optimization (GO) method is usually selected [8,9]. There are deterministic and stochastic methods for GO [25,26]. Stochastic methods have been showing to be robust alternatives to the

deterministic and exact methods. The stochastic alternatives to solve GO problems are usually general-purpose, population-based and are normally referred to as evolutionary algorithms since they are motivated by biological evolution [27]. The genetic algorithm [28,29] is the most known evolutionary strategy, although other swarm intelligence based algorithms [30,31,32,33] are common in the literature. A recent strategy successfully used in biological and engineering applications is the FA [34].

The FA [12] is a stochastic approach for GO that is inspired by the collective behavior of fireflies, specifically in how they attract each other. Several variants of FA have been designed so far, but the main simple rules that define the firefly movement are the following [35]:

- all fireflies are unisex, meaning that any firefly will be attracted to other fireflies regardless of their sex;
- the brightness of a firefly is determined by the value of the objective function;
- attractiveness of a firefly is proportional to their brightness; thus for any two fireflies, the less brighter will move towards the brighter one; the attractiveness is directly proportional to the brightness but decreases with distance.

Hereafter, to represent the position of the firefly i the following vector notation is used $p^i = (p_1^i, \ldots, p_5^i)$. Initially, the positions of the population of m fireflies are randomly generated in the search space defined by the bound constraints on the decision variables, as follows

$$p_s^i = p_s^{\mathrm{L}} + rand^i(p_s^{\mathrm{U}} - p_s^{\mathrm{L}}), \;\; \text{for} \;\; s = 1, \ldots, 5 \;\; \text{and} \;\; i = 1, \ldots, m \tag{6}$$

where i stands for the firefly i in the population of m fireflies, s represents the component of the vector, p^{L} and p^{U} are the vectors of lower and upper bounds on the decision variables, and $rand^i$ is a vector of random numbers from a uniform distribution on the interval $[0, 1]$.

The movement of firefly i, in direction to a brighter firefly j is defined by

$$p^i = p^i + \beta(p^j - p^i) + \alpha \; rand^{i,j} \tag{7}$$

where $0 \leq \alpha \leq 1$ is the randomness parameter and $rand^{i,j}$ is a vector of uniformly distributed numbers in the range $[-1, 1]$, and β is given by:

$$\beta = \beta_0 e^{-\gamma ||p^i - p^j||^v} \;\; \text{for} \;\; v \geq 1, \tag{8}$$

being β_0 the attraction when the distance is zero. The parameter $\gamma \in [0, \infty)$ characterizes the variation of the attractiveness and it is crucial to determine the algorithm convergence speed. When $\gamma \to 0$, attractiveness is constant which means that the brightness of the firefly can be seen from any position of the space. On the other hand, when $\gamma \to \infty$ the attractiveness is almost zero and each firefly moves randomly. Note that, in the movement described by (7), the second term is due to the attraction and the third term is due to randomization.

4 Numerical Experiments

In this section, the practical performance of the metaheuristic FA for solving the BCNLP problem that emerges from the implementation of the direct sequential procedure to the α-pinene isomerization problem, is analyzed. The mathematical model was coded in the Matlab programming language and the computational application of the direct sequential procedure with the FA was developed in Matlab programming environment. The computational tests were performed on a PC with a 2.7 GHz Core i7-4600U and 8 Gb of memory.

4.1 FA Parameter Tuning

First, an experimental study was conducted to tune some of the FA parameters, in the context of the α-pinene problem. The parameter in the randomization term is made to depend on the allowed maximum number of iterations and decreases from an initial value, α, to a near zero value. Four values are tested, ranging from the largest possible to a rather small one, $\alpha = \{1, 0.5, 0.25, 0.05\}$, being 0.25 the most common value used in the literature. The parameter γ is kept fixed over the iterative process. To analyze the effect of the attractiveness in the algorithm, four moderate values have been selected: $\gamma = \{2, 1, 0.5, 0.25\}$. Only eight combinations of α and γ have been tested in these experiments. A population of 20 points is used and the problem is solved 10 times with independent randomly generated initial populations. The iterative FA terminates when the absolute difference between the obtained solution, at the current iteration, and the best known optimal solution is below a small tolerance, for instance 1.0e-4, or when a maximum of 500 iterations is reached. The lower and upper bounds for the decision variables are p_i^{L}=0 and p_i^{U}=5.0e-4 for $i = 1, \ldots, 5$. The best known solution for this problem is $p_1^* = 5.9256\text{e-}5$, $p_2^* = 2.9632\text{e-}5$, $p_3^* = 2.0450\text{e-}5$, $p_4^* = 2.7473\text{e-}4$, $p_5^* = 4.0073\text{e-}5$, with the optimum objective function value $J(p^*) = 19.872$, as reported in [36].

Table 2 contains the results produced by the proposed algorithm and correspond to the best function value, J_{best}, the average of the function values, over the 10 runs, J_{avg}, the median of the function values, J_{med}, and the average number of iterations, $N.It_{\mathrm{avg}}$. From the table, it can be concluded that combining $\alpha = 0.25$ with $\gamma = 1$ gives the best results in terms of best, average and median J values, and a lower average number of iterations. The run that produced the best value $J_{\mathrm{best}} = 19.8772$ required 294 iterations and 5860 function evaluations. From the table, it is possible to conclude that the two best sets of results correspond to a small value of α, meaning that the randomness characteristic of the FA is somehow controlled, and to a moderate value of γ (0.5 and 1) which promotes the local search ability of the FA. We note that a very small value of α (for example, 0.05) is harmful since the algorithm is not capable of exploring the search space for the region where the global solution lies.

Figure 3 shows the evolution of the objective function along the iterations, relative to the best and worst runs for the case where $\alpha = 0.25$ and $\gamma = 1$. For comparative purposes, the evolution of the objective function of the worst run,

for the case where $\alpha = 0.5$ and $\gamma = 0.5$ is also shown. As it can be seen the value $\alpha = 0.25$ produces a significant decrease in J during the early iterations.

Table 2. Results obtained by FA, for different combinations of α and γ.

α	γ	J_{best}	J_{avg}	J_{med}	$N.It_{\text{avg}}$
1	1	20.0334	30.2361	28.2366	500
0.5	1	20.4490	25.7785	25.5231	500
0.5	0.5	19.9366	67.8867	29.4598	500
0.25	2	19.8819	44.2389	30.5656	490
0.25	1	19.8772	25.6777	20.7662	459
0.25	0.5	19.8970	27.0022	24.2394	500
0.25	0.25	19.9029	356.7275	180.7635	500
0.05	1	260.6788	979.0378	657.9491	500

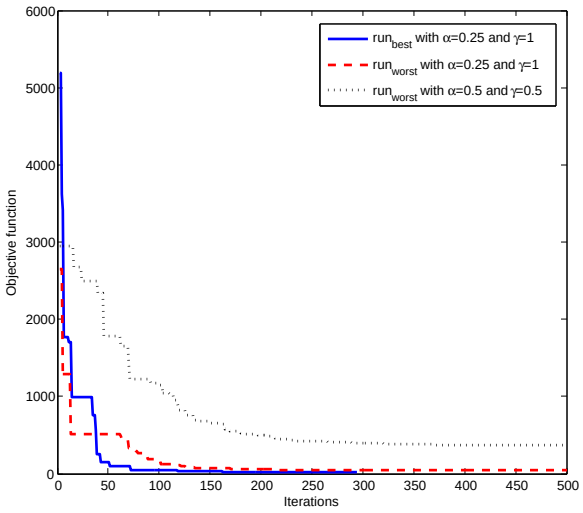

Fig. 3. Evolution of the objective function from the best and worst runs, when $\alpha = 0.25$ and $\gamma = 1$, and from the worst run, when $\alpha = 0.5$ and $\gamma = 0.5$.

4.2 Comparative Results

The second experiment aims to compare the results produced by the FA with those obtained by the `fmincon` function from the Matlab Optimization Toolbox, as well as with those achieved in [14,18,36,37,38]. Therefore, Table 3 contains

a summary of the results – the best objective function value, the best parameter values, and the number of iterations, $N.It$, and the number of function evaluations, $N.f.ev$, of the reported solution – obtained by:

- FA (in the context of the direct sequential approach by invoking the function `ode15s` for the numerical integration of the ODEs);
- `fmincon` (in the context of the direct sequential approach with `ode15s` for the numerical integration of the ODEs);
- the hybrid scatter search methodologies proposed in [36,18] – therein denoted by SSm – (in the context of the direct sequential approach);
- the metaheuristic EM, implemented with a random local search procedure, in [37] (in the context of the direct sequential approach, with `ode45` for the numerical integration of the ODEs);
- the deterministic global optimization algorithm based on outer approximation proposed in [14] (in the context of orthogonal collocation on finite elements in $[0, 36420]$);
- the deterministic gradient-based solver 'filterSQP' (a sequential quadratic programming based on the filter set methodology) tested in [38] (in the context of 3-stage collocation method with uniform partition of 100 subintervals in $[0, 36420]$).

Table 3. Optimal objective function value and parameters: FA *vs* `fmincon` and the results in [14,18,36,37,38].

	solution	parameter vector	$N.It$	$N.f.ev$
FA	19.877	(5.928e-5, 2.962e-5, 2.041e-5, 2.764e-4, 4.046e-5)	294	5860
`fmincon`	19.929	(5.911e-5, 2.968e-5, 2.070e-5, 2.745e-4, 4.001e-5)	24	217
in [36]	19.872	n.a.	n.a.	1163[a]
in [18]	19.87	(5.926e-5, 2.963e-5, 2.047e-5, 2.745e-4, 3.998e-5)	9518	(*122*)[b]
in [37]	19.874	(5.926e-5, 2.963e-5, 2.047e-5, 2.743e-4, 3.991e-5)	147	3824
in [14]	19.87	n.a.	2	(*8916*)[b]
in [38]	19.8721	n.a.	n.a.	(*28*)[b]

n.a. means not available.
[a] Average value over ten independent runs.
[b] CPU time in seconds.

The `fmincon` function implements a local gradient-based method and requires an initial approximation provided by the user. The sequential quadratic programming based on a quasi-Newton update formula to approximate the second derivatives of the objective function is used, and `fmincon` was invoked with its default settings. It is common knowledge that traditional local solvers may fail to reach a global optimal solution unless a good initial approximation is provided by the user. Thus, the initial approximation $p_i = 0, i = 1, \ldots, 5$ is used. The `fmincon` function achieved the minimum $J = 19.929$ after 24 iterations and

217 function evaluations. The best solution reported by FA $J_{\text{best}} = 19.8772$ (see Table 2) is close to the best known optimal solution, with a relative error of 0.026%, while the solution obtained with `fmincon` has a relative error of 0.29%. This is a confirmation that FA makes a comprehensive exploration of the search space, without requiring any good initial approximation. The scatter search algorithms proposed in [18,36] activate a gradient-based local search procedure (the `fmincon` function from Matlab) that is carried out from different vectors as initial points to accelerate convergence to the solution. In [36], the best value $J_{\text{best}} = 19.872$, the average value $J_{\text{avg}} = 24.747$, the worst value $J_{\text{wor}} = 68.617$ and the average number of evaluations $N.f.ev_{\text{avg}} = 1163$ (over ten executed runs) are reported. In [18], the initial approximation $p_i = 0.5, i = 1, \ldots, 5$ is used and the optimal value $J = 19.87$ is reached after 9518 iterations and 122 seconds. In [37], the best value $J_{\text{best}} = 19.874$, the median value $J_{\text{med}} = 20.325$, the worst value $J_{\text{wor}} = 21.806$, and the average number of iteration $N.It_{\text{avg}} = 147$ and of function evaluations $N.f.ev_{\text{avg}} = 4070$ (over ten executed runs) are reported. The outer approximation based algorithm presented in [14] converges to $J = 19.87$. Although only two iterations of the outer approximation strategy are required to reach the solution, the therein used approach may lead to computational burdens and the process is a very costly one (see details in Table 3). The problem at hand is also part of the well-known Large-Scale Constrained Optimization Problem Set (COPS) and was solved by a collocation method and several NLP solvers, for comparative purposes, in [38]. The minimum J value found by the well-known 'filterSQP' was 19.8721 after 28 seconds. It can be concluded from the results shown in the table that the direct sequential approach based on the FA is effective in reaching the optimal solution. Convergence to the solution may be accelerated and solution quality may be improved by incorporating a local search technique into FA.

Figure 4 shows the evolution over time of the concentrations of reactant and the four products for the α-pinene isomerization problem, considering the optimal values of the parameters obtained by the FA (shown in Table 3) – represented by full lines. The figure also shows the experimental data represented by points alone, as reported in [19]. As it can be seen the obtained optimal values for the parameters allow the model to reproduce almost exactly the experimental data. On the other hand, Fig. 5 shows the experimental data (represented by the points alone) and the values predicted by the model (represented by full lines) for the same time instants, considering the optimum values for the parameters produced by `fmincon` (as reported in Table 3).

5 Conclusions

We have shown that the parameter estimation problem on the α-pinene isomerization model can be easily solved by a direct sequential approach which transcribes the DO problem in a small finite-dimensional BCNLP problem. As a consequence, a stochastic global optimization technique can be applied to locate a global minimum of the goodness of fit objective function subject to bound con-

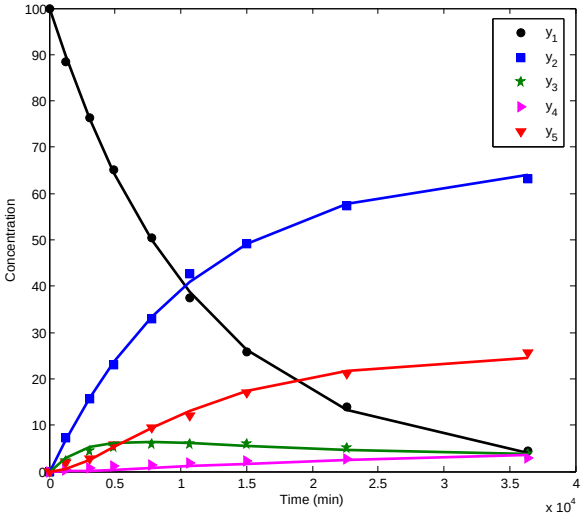

Fig. 4. Concentrations estimated by the model (full lines) and experimentally observed values (points alone) in the isomerization of α-pinene, based on the parameters produced by FA.

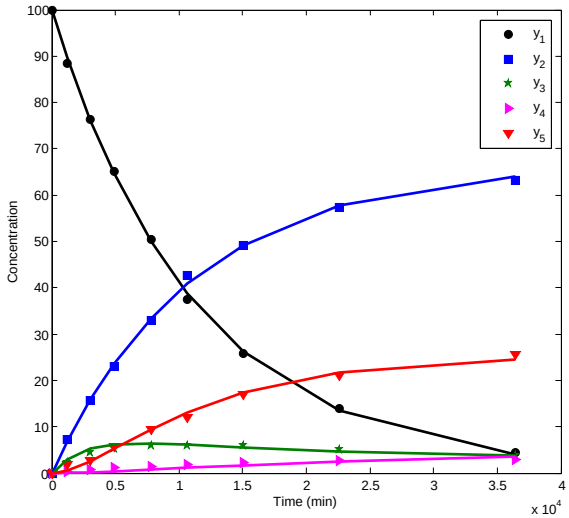

Fig. 5. Concentrations estimated by the model (full lines) and experimental values (points alone) in the α-pinene isomerization, based on the parameters obtained by `fmincon`.

straints on the values of the parameters. The simple to understand and easy to implement metaheuristic FA has shown to be a reliable method to locate a global optimal solution of this challenging DO problem. Some preliminary experiments

were carried out in order to tune the parameters α and γ of the FA, and it is shown that this issue is crucial to the convergence features of the algorithm. The implementation of FA was effective in a way that the produced results are comparable to those of previously used metaheuristics and outperform the results obtained by `fmincon` function from Matlab. The obtained solution is very close to the best known solution in the literature and allows the model to reproduce almost exactly the experimental data.

Future developments will be directed to the FA enhancing strategy by using an efficient but simple random local search procedure so that convergence could be accelerated.

Acknowledgments. The authors wish to thank two anonymous referees for their comments and suggestions. This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT - Fundação para a Ciência e Tecnologia, within the projects UID/CEC/00319/2013 and UID/MAT/00013/2013.

References

1. Banga, J.R., Balsa-Canto, E., Moles, C.G., Alonso, A.A.: Improving food processing using modern optimization methods. Trends Food Sci. Tech. 14(4), 131–144, (2003)
2. Grossmann, I.E. (ed.): Global Optimization in Engineering Design, Nonconvex Optimization and Its Applications vol. 9, Springer Science & Business Media (1996)
3. Vanrolleghem, P.A., Dochain, D.: Bioprocess model identification. In: Van Impe, J.F.M., Vanrolleghem, P.A., Iserentant, D.M. (eds.), Advanced Instrumentation, Data Interpretation, and Control of Biotechnological Processes, pp. 251–318, Kluwer Academic Publ. (1998)
4. Rodríguez-Fernández, M.: Modelado e Identificación de Bioprocesos. PhD thesis, University of Vigo, Spain (2006)
5. Rodriguez-Fernandez, M., Mendes, P., Banga, J.R.: A hybrid approach for efficient and robust parameter estimation in biochemical pathways. Biosystems, 83(2-3), 248–265, (2006)
6. Gill, P.E., Murray, W., Saunders, M.A., and Wright, M.H. : Constrained Nonlinear Programming. In: Handbooks in Operations Research and Management Science, vol. 1, Optimization, pp. 171–210 (1989)
7. Banga, J. R., Balsa-Canto, E., Moles, C. G., and Alonso, A. A.: Dynamic optimization of bioreactors: a review. In: Proceedings-Indian National Science Academic, Part A, 69(3/4), 257–266 (2003)
8. Esposito, W.R. and Floudas, C.A.: Global optimization for the parameter estimation of differential-algebraic systems. Ind. Eng. Chem. Res. 39(5), 1291–1310 (2000)
9. Moles, C.G., Mendes, P., Banga, J.R.: Parameter estimation in biochemical pathways: a comparison of global optimization methods. Genome Res. 13(11), 2467–2474 (2003)
10. Moles, C.G., Gutierrez, G., Alonso, A.A., Banga, J.R.: Integrated process design and control via global optimization: a wastewater treatment plant case study. Chem. Eng. Res. Des. 81(5), 507–517 (2003)

11. Eberhart, R.C., Kennedy, J.: A new optimizer using particle swarm theory. In: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, vol. 1, pp. 39–43, New York, NY, (1995)
12. Yang, X.-S.: Firefly algorithms for multimodal optimization. In: Steinhöfel, K. (ed.), Stochastic Algorithms: Foundations and Applications, International Symposium SAGA 2001, pp. 169–178. Springer (2009)
13. Lin, Y. and Stadtherr, M.A.: Deterministic global optimization for parameter estimation of dynamic systems. Ind. Eng. Chem. Res. 45(25), 8438–8448 (2006)
14. Miro, A., Pozo, C., Guillén-Gosálbez, G., Egea, J.A., Jiménez L.: Deterministic global optimization algorithm based on outer approximation for the parameter estimation of nonlinear dynamic biological systems. BMC Bioinformatics, 13, 90 (2012)
15. Polisetty, P.K., Voit, E.O., Gatzke, E.P.: Identification of metabolic system parameters using global optimization methods. Theor. Biol. Med. Model. 3, 4 (2006)
16. Balsa-Canto, E., Peifer, M., Banga, J.R., Timmer, J., Fleck, C.: Hybrid optimization method with general switching strategy for parameter estimation. BMC Syst. Biol. 2, 26 (2008)
17. Box, G., Hunter, W., MacGregor, J., Erjavec, J.: Some problems associated with the analysis of multiresponse data. Technometrics, 15(1), 33–51 (1973)
18. Rodriguez-Fernandez, M., Egea, J.A., Banga, J.R.: Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. BMC Bioinformatics, 7, 483, (2006)
19. Fuguitt, R.E., Hawkins, J.E.: Rate of the thermal isomerization of α-pinene in the liquid phase. J. Am. Chem. Soc. 69(2), 319–322 (1947)
20. Hunter, W., McGregor, J.: The estimation of common parameters from several responses: Some actual examples. Unpublished Report, Department of Statistics. University of Winsconsin (1967)
21. Srinivasan, B., Palanki, S., Bonvin, D.: Dynamic optimization of batch processes: I. Characterization of the nominal solution. Comput. Chem. Eng. 27(1), 1–26 (2003)
22. Banga, J.R., Balsa-Canto, E., Moles, C.G., Alonso, A.A.: Dynamic optimization of bioprocesses: Efficient and robust numerical strategies. J. Biotechnology, 117, 407–419 (2005)
23. Schlegel, M., Marquardt, W.: Direct sequential dynamic optimization with automatic switching structure detection. In: Shah, S.L., MacGregor, J. (eds.), Dynamics and Control of Process Systems 2004 (DYCOPS-7), vol. 1, pp. 419–424, Elsevier IFAC Publ. (2005)
24. Biegler, L.T.: Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes. MOS-SIAM Series on Optimization (2010)
25. Zhigljavsky, A., Žilinskas, A.: Stochastic Global Optimization. Springer Optimization and Its Applications, Springer (2008)
26. Floudas, C.A., Akrotirianakis, I.G., Caratzoulas, S., Meyer, C.A., Kallrath, J.: Global optimization in the 21st century: Advances and challenges. Comput. Chem. Eng. 29, 1185–1202 (2005)
27. Tsai, K.-Y., Wang, F.-S.: Evolutionary optimization with data collocation for reverse engineering of biological networks. Bioinformatics, 21(7), 1180–1188 (2005)
28. Goldberg, D.E., Korb, B., Deb, K.: Messy genetic algorithms: Motivation, analysis, and first results. Complex Systems, 3(5), 493–530 (1989)
29. Holland, J.H.: Genetic algorithms. Scientific American, 267(1), 66–72 (1992)
30. Bonabeau, E., Dorigo, M., Theraulaz, G.: Swarm Intelligence: From Natural to Artificial Systems. Santa Fe Institute Studies on the Sciences of Complexity, Oxford University Press (1999)

31. Kennedy, J., Eberhart, R.C.: Swarm Intelligence. Morgan Kaufmann Publishers Inc., San Francisco, USA (2001)
32. Millonas, M.M.: Swarms, phase transitions, and collective intelligence. In: Langton, C.G. (ed.), Artificial Life III, vol. XVII, pp. 417–445, Addison-Wesley, Reading, MA (1994)
33. Yang, X.-S.: Nature-Inspired Metaheuristic Algorithms. 2nd edition, Luniver Press (2010)
34. Yang, X.-S.: Biology-derived algorithms in engineering optimization. In: Olariu, S., Zomaya, A. Y. (eds.), Handbook of Bioinspired Algorithms and Applications, pp. 589–600, Chapman & Hall (2005)
35. Yang, X.-S.: Firefly algorithm, stochastic test functions and design optimisation. Int. J. Bio-Inspired Computation, 2(2), 78–84 (2010)
36. Egea, J.A., Rodríguez-Fernández, M., Banga, J.R., Marti, R.: Scatter search for chemical and bioprocess optimization. J. Glob. Optim. 37(3), 481–503 (2007)
37. Rocha, A.M.A.C., Silva, A., Rocha, J.G.: A new competitive implementation of the electromagnetism-like algorithm for global optimization. In: Gervasi, O. et al. (eds.), ICCSA 2015, Part II, Lect. Notes Comput. Sc., vol. 9156, pp. 506–521, Springer (2015)
38. Dolan, E.D., Moré, J.J., Munson, T.S.: Benchmarking Optimization Software with COPS 3.0. Argonne National Laboratory Technical Report ANL/MCS-TM-273 (2004)