Evolutionary Algorithms for Parameter Estimation of Metabolic Systems

Anastasia Slustikova Lebedik and Ivan Zelinka

Abstract. For many years, computational tools have been widely applied to study such complex systems as metabolic networks. One of the principal questions in modeling of metabolic systems is the parameter estimation of model, which is related to a nonlinear programming problem. Two types of evolutionary algorithms, Differential Evolution and Self-Organizing Migrating Algorithm, are applied to the well-studied metabolic system, the urea cycle of the mammalian hepatocyte. The algorithms provide an effective approach in parameters identification of the model.

1 Introduction

Metabolic systems are highly non-linear with complex structure and dynamics. The complexity of interactions between components of metabolic system makes the prediction of the system behavior extremely challenging [1], [2]. To overcome this challenge, many researchers use computational and mathematical modeling methods.

From mathematical point of view, metabolic systems can be described in terms of ordinary differential equations (ODEs). In particular, the formulation of the system of ODEs for metabolic systems requires knowledge of biochemical reaction mechanisms. The dynamic behavior of the system is characterized by a set of parameters such as kinetic constants or rate constants. These parameters are often unknown due to highly complicated or even impossible experimental determination. In addition, most parameters defined in experiments *in vitro* are different

Anastasia Slustikova Lebedik
Tomas Bata University in Zlin, Faculty of Applied Informatics, T.G. Masaryka 5555,
760 01 Zlin, Czech Republic
e-mail: lebedik@fai.utb.cz

Ivan Zelinka
VŠB-Technical University of Ostrava, Faculty of Electrical Engineering and
Computer Science, Department of Computer Science, 17. listopadu 15,
708 33 Ostrava-Poruba, Czech Republic
e-mail: ivan.zelinka@vsb.cz

I. Zelinka et al. (Eds.): *Nostradamus 2013: Prediction, Model. & Analysis*, AISC 210, pp. 201–209.
DOI: 10.1007/978-3-319-00542-3_21

from parameters measured *in vivo* [3]. Therefore, the estimation of the model parameters is required for prediction of the system dynamics.

Due to a large number of reactions, non-linear interactions between different metabolites, enzymes and other components of the system, the parameters estimation of metabolic systems can be formulated as non-linear programming (NLP) problem [4], [5]. A common approach for solving such a problem is the application of optimization methods. The optimization task is to minimize the difference between experimentally measured and the simulated data.

Different techniques are used to find the model parameters that make the model best fitting to experimental measurements. Application of local optimization algorithms is usually limited because of convergence to local optima [6]. Whereas using global optimization methods have been successfully applied to the parameter estimation problem. Among global optimization techniques, the use of evolutionary algorithms (EAs) should be highlighted. EAs are known as an efficient approach that can cope with large-scale systems. Recent studies have successfully applied EAs to parameter optimization problem [3], [5], [6] and others.

In present study, we have applied two modern evolutionary techniques to parameter estimation of the well-studied metabolic system, the urea cycle of the mammalian hepatocyte.

In our investigation, we consider three main questions:

1. Are the algorithms capable of precise parameter estimation of the urea cycle model?
2. Does the performance of studied evolutionary techniques depend on various algorithms settings?
3. Which algorithm performs best on parameter estimation of the metabolic system?

2 Experiment Design

In this research, the urea cycle is chosen for parameter optimization using evolutionary algorithms. Two modern efficient evolutionary techniques are applied to an optimization task which is described below.

2.1 The Urea Cycle Model

Our simulations were based on the mathematical model of the urea cycle of the mammalian hepatocyte described by P. W. Kuchel and P. J. Mulquiney [7]. Variation of different kinetic parameters of the enzymes of the urea cycle can effect metabolite concentration.

The model of urea cycle was developed to investigate dependence of metabolite concentrations on various kinetic parameters of the enzymes. The model includes four enzyme reaction schemes: arginase, ornithine carbomoyl transferase, argininosuccinate lyase, and argininosuccinate synthetase (see Fig. 1.). The metabolic model was formed by 12 ODEs and a set of kinetic parameters.

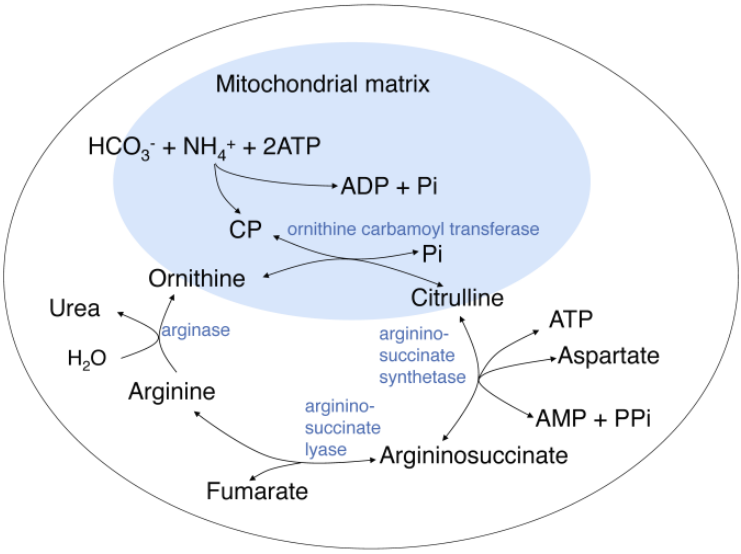

Fig. 1 The urea cycle in mammalian hepatocyte [7]

The detail description of the system can be found in the original source [7]. Overall, the number of kinetic parameters was 45. Due to the large number of unknown parameters, the parameter estimation of this problem is related to NLP problem.

2.2 Cost Function

Generally, parameter estimation of nonlinear systems is formulated as a task of minimization of cost function. The cost function in our research was stated as the sum of differences between experimentally measured and simulated data.

In our case, the experimentally measured data were replaced by simulated data using the true values of the model parameters obtained from [7]. Initial concentration of the studied metabolites: ornithine, citrulline, arginine and urea, were taken from the same source as the true values of the parameters.

2.3 Used Algorithms

Two types of evolutionary algorithms: Differential Evolution [8] and Self-Organizing Migrating Algorithm [9] have been considered in the experiments. The choice of DE was based on its published performance. And SOMA is a new and perspective global optimization technique.

DE is a floating-point encoded evolutionary algorithm for global optimization introduced by Storn and Price, see [8]. The initial population is randomly selected. New population members are generated using recombination and mutation. The

performance of DE depends on the choice of the mutation and crossover strategies and control parameters such as the population size, crossover rate and the scale factor. The main advantages of DE are finding true global minimum regardless of the initial parameter values, fast convergence and using few control parameters.

The second used algorithm, called SOMA (see [9]), was used in this research together with DE. The SOMA algorithm is based on a cooperative-competitive principle. The main difference SOMA from other evolutionary techniques is that the new population is formed by principle, which is based on the social behavior of cooperating individuals. Moreover, unlike other evolutionary algorithms, SOMA uses a special parameter (PRT) instead of mutation. The crossover operator can be thought as the movement of an individual.

2.4 Algorithms Setting and Used Hardware

For the DE approach, we have firstly studied the impact of population size on the algorithm performance. Then, we have investigated the influence of factor F, which controls the amplification of the differential variation. The above mentioned settings with the minimum cost function value have been used for identification the best value of CR, the crossover constant. The details of the experiments are described in the section Results.

To study the performance of SOMA algorithm, we have varied only population size (PopSize).

The minimum of cost function value has been used as a quality measure of every set of algorithm settings.

The experiments were conducted using *Mathematica* 7. Each experiment was repeated 40 times. We have used the DERand1Bin version of DE and the AllToOne version of SOMA.

All calculations have been done using grid computer that includes 16 XServers, each 2x2 GHz Intel Xeon, 1 GB RAM, 80 GB HD i.e. 64 CPUs.

3 Results

Two evolutionary computation techniques: DE and SOMA were applied to well-studied metabolic system, the urea cycle of the mammalian hepatocyte. The model includes four enzyme reaction schemes: arginase, ornithine carbomoyl transferase, argininosuccinate lyase, and argininosuccinate synthetase.

The efficiency of particular method was judged based on cost function value.

All optimization techniques yield meaningful results. Each algorithm is capable of finding the model parameters. Performance of DE and SOMA is presented in Figure 2.

DE gives the best result with F=0.8, CR=0.6 and population size 900. The minimum of cost function value with these settings is 2.13×10^{-4}. In contrast to DE, SOMA reaches the best cost function of 5.61×10^{-5} with PopSize=135. However, in

order to compare the performance of these two algorithms, we also take into account number of cost function evaluations. Hence, there have been chosen the results of SOMA simulations with PopSize=90. In this case, the cost function value is 8.43×10^{-5}, which is slightly higher but still lower in comparison with the best DE result. The implemented experimental setting in the above mentioned simulations are presented in Table 1.

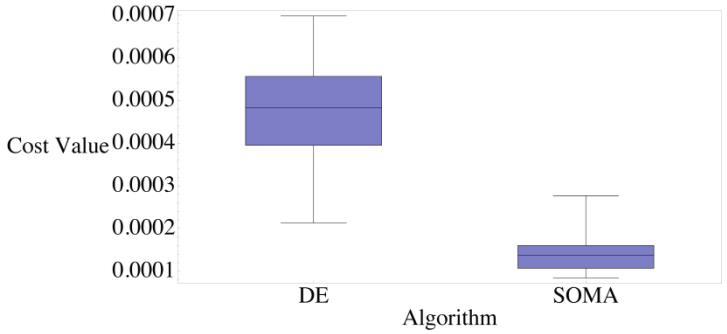

Fig. 2 Comparison of the optimization algorithms applied to the urea cycle model

Table 1 Settings for the algorithms

DE settings		SOMA settings	
NP	900	PathLength	3
F	0.8	Step	0.11
CR	0.6	PRT	0.1
Generations	150	PopSize	90
		Migrations	50
		MinDiv	-0.1

3.1 DE Experiments

To find the best settings for DE algorithm, we vary population size, the F value and CR. The minimum of cost function value is used as a quality measure of every set of algorithm settings.

Figure 3 depicts dependence of cost function value on various population sizes. To investigate the impact of population size, we apply a population size of 90, 450 and 900. These settings are equal to 2D, 10D and 20D, where D is a number of cost function arguments. In our case, it is 45.

We limit the investigation to only 3 sets of population size because of the execution time, which depends on the dimension of the problem. The dimension of the studied problem is high (45 parameters). Each experiment is repeated 40 times.

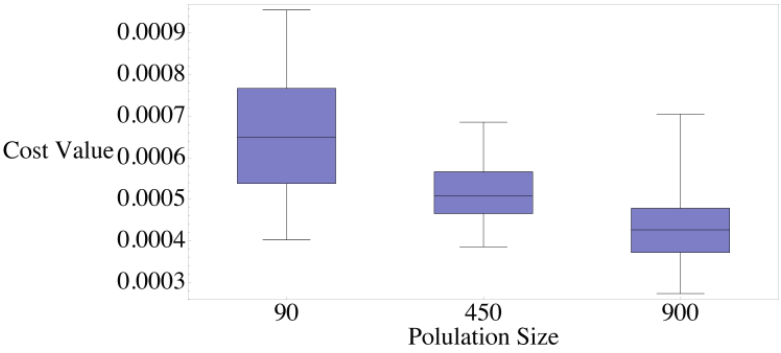

Fig. 3 The impact of population size on the cost function value for DE

The boxplots show that a population size of 900 reaches the best minimum result. The average cost function value decreases with increasing population size. However, all results give very low cost function value.

We continue to study the DE performance by varying the values for F (see Fig.4.) and CR (see Fig.5.). The influence of the F is tested using three F values: 0.1, 0.5 and 0.8. The CR value is set on 0.5. The DE algorithms yields the best results for F=0.8. Therefore, the best settings F=0.8 and population size of 900 are used for the next investigation.

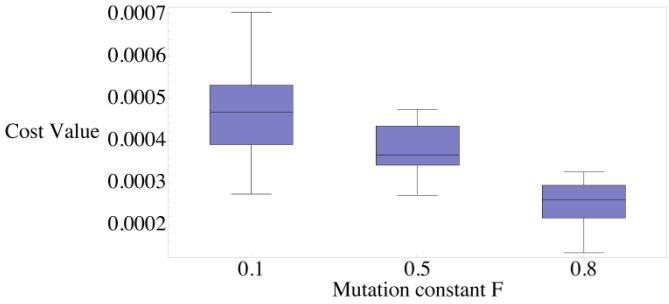

Fig. 4 The influence of mutation constant F on the cost function value for DE

To find the most successful combination of the algorithm settings, we vary CR from 0.1 to 0.9. Similarly to the above mentioned experiments, each calculation is repeated 40 times. Figure 5 depicts that CR=0.6 yields the best minimum result. It should be noticed that DE with all values of CR reaches meaningful results with the cost function value from 2.13×10^{-4} to 9.48×10^{-4}.

Figure 6 shows simulation of the system dynamics using predicted parameters (dashed) together with original parameters (solid). There are time courses of 4 main metabolites concentrations in the urea cycle simulation: ornithine, citrulline, arginine and urea. The figure depicts the best result of DE algorithm with F=0.8, CR=0.6 and population size 900.

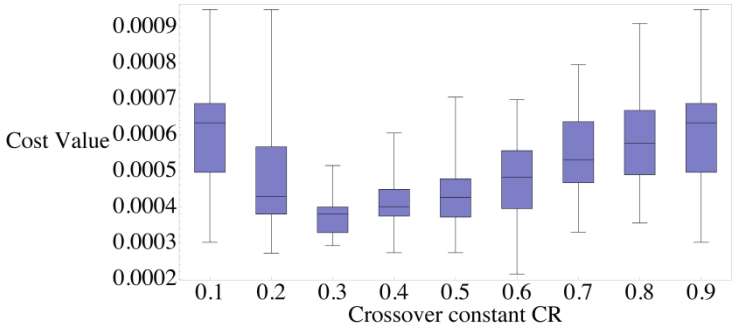

Fig. 5 The influence of crossover constant CR on the cost function value for DE

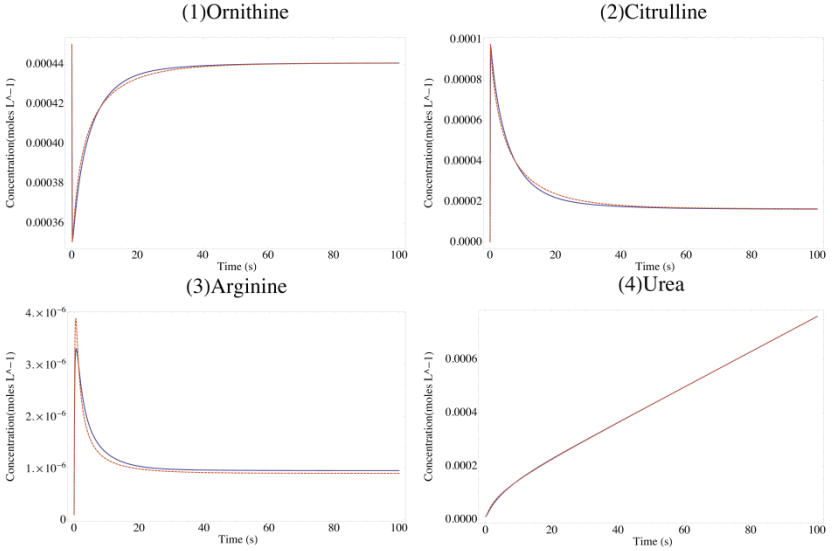

Fig. 6 The time courses of (1) ornithine, (2) citrulline, (3) arginine and (4) urea where predicted behavior by DE is dashed and original is solid

The DE algorithm performs very well. It is obvious that parameters of the model are predicted precisely.

3.2 SOMA Experiments

To study the performance of SOMA algorithm, we vary only population size (PopSize). The minimum of cost function value is used as a quality measure of every set of algorithm settings.

Similarly to DE, the study is limited to 3 sets of settings with population size of 45, 90 and 135, which equal 1D, 2D and 3D. The choice of population sizes for

SOMA algorithm is based on recommendations in [9]. The number of repetitions is again 40. The results of the experiments are shown in Figure 7.

The boxplots show that varying PopSize has similar impact on estimation process as in case of DE - the higher population size, the lower cost function value. SOMA yields the minimum of cost function value5.61×10^{-5} with PopSize=135. The worst result of SOMA algorithm is 2.77×10^{-4}, which is slightly higher than the best result of DE algorithm 2.13×10^{-4}.

Figure 8 depicts the time courses of 4 urea cycle metabolites with predicted and original parameters. The behavior of the system is predicted precisely.

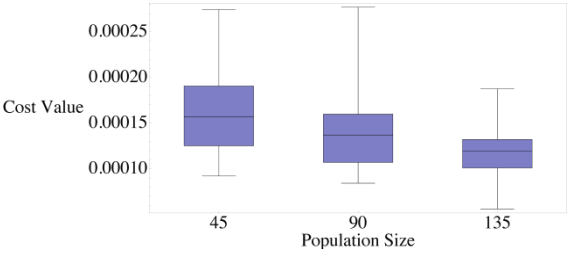

Fig. 7 The impact of population size (PopSize) on the cost function value for SOMA

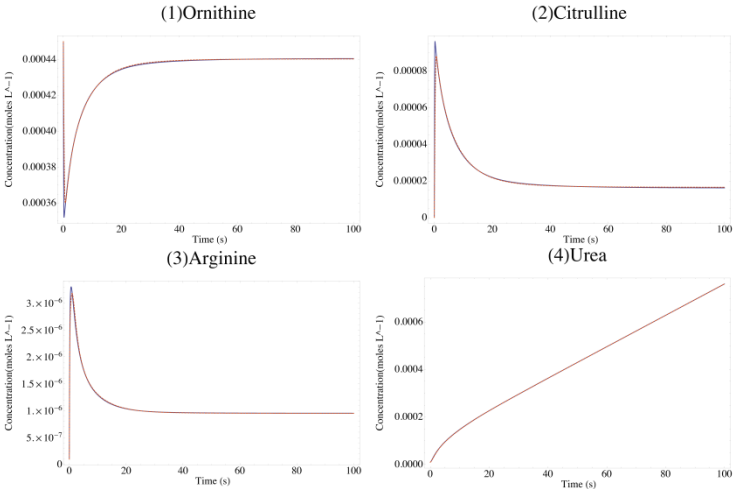

Fig. 8 The time courses of (1) ornithine, (2) citrulline, (3) arginine and (4) urea where predicted behavior by SOMA is dashed and original is solid

4 Conclusions

All applied optimization techniques yielded meaningful results. Each algorithm was capable of precise parameter estimation of the urea cycle model.

Varying algorithms settings could improve the algorithm performance. In DE simulations the influence of mutation constant F and crossover constant CR was significant. The best combination of these constants for the urea cycle simulations is F=0.8 and CR=0.6.In both cases DE and SOMA, increasing population size gave significantly better results, see Fig.3 and Fig.7. On the other hand, it required large computational effort. Considering computational time, the most time-consuming calculations were observed in SOMA simulations. However, it should be noted that SOMA provided the best performance in estimating parameters.

Taking into account that all algorithms performed well, and the fastest was DE, it is reasonable to apply DE in experiments with limited computational time.

Acknowledgments. This work was supported by grant No. IGA/FAI/2013/005 of the Internal Grant Agency of Tomas Bata University in Zlin.

References

1. Sun, J., Garibaldi, J.M., Hodgman, C.: Parameter Estimation Using Metaheuristics in Systems Biology: A Comprehensive Review. IEEE/ACM Transactions On Computational Biology and Bioinformatics 9(1), 185–202 (2012)
2. Mendes, P., Kell, D.B.: Non-Linear Optimization of Biochemical Pathways: Applications to Metabolic Engineering and Parameter Estimation. Bioinformatics 14(1), 869–883 (1998)
3. Dräger, A., Kronfeld, M., Ziller, M.J., Supper, J., Planatscher, H., Magnus, J.B., Oldiges, M., Kohlbacher, O., Zell, A.: Modeling metabolic networks in C. glutamicum: a comparison of rate laws in combination with various parameter optimization strategies. BMC Systems Biology 3(5) (2009)
4. Moles, C.G., Mendes, P., Banga, J.R.: Parameter Estimation in Biochemical Pathways: A Comparison of Global Optimization Methods. Genome Research 13, 2467–2474 (2003)
5. Polisetty, P.K., Voit, E.O., Gatzke, E.P.: Identification of Metabolic System Parameters Using Global Optimization Methods. Theoretical Biology and Medical Modelling 3(4), 1–15 (2006)
6. Rodriguez-Fernandez, M., Egea, J.A., Banga, J.R.: Novel Metaheuristic for Parameter Estimation in Nonlinear Dynamic Biological Systems. BMC Bioinformatics 7, 483–501 (2006)
7. Mulquiney, P.J., Kuchel, P.W.: Modeling metabolism with Mathematica. CRC Press (2003)
8. Price, K., Storn, R.M., Lampinen, J.A.: Differential Evolution: A Practical Approach to Global Optimization. Natural Computing. Springer (1995)
9. Zelinka, I.: SOMA – Self Organizing Migrating Algorithm. In: Babu, B.V., Onwubolu, G. (eds.) New Optimization Techniques in Engineering, vol. 33, ch. 7. Springer (2004)