

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# An Improved Algorithm for Optimising the Production of Biochemical Systems

*Mohd Arfian Ismail, Vitaliy Mezhuyev,  
Mohd Saberi Mohamad, Shahreen Kasim  
and Ashraf Osman Ibrahim*

## Abstract

This chapter presents an improved method for constrained optimisation of biochemical systems production. The aim of the proposed method is to maximise its production and, at the same time, to minimise the total amount of chemical concentrations involved in producing the best production. The proposed method models biochemical systems with ordinary differential equations. The optimisation process became complex for the large size of biochemical systems that contain many chemicals. In addition, several constraints as the steady-state constraint and the constraint of chemical concentrations also contributed to the computational complexity and difficulty in the optimisation process. This chapter considers the biochemical systems as a nonlinear equations system. To solve the nonlinear equations system, the Newton method was applied. Then, both genetic algorithm and cooperative co-evolutionary algorithm were applied to fine-tune the components in the biochemical systems to maximise the production and minimise the total amount of chemical concentrations involved. Two biochemical systems were used, namely the ethanol production in the *Saccharomyces cerevisiae* pathway and the tryptophan production in the *Escherichia coli* pathway. In evaluating the performance of the proposed method, several comparisons with other works were performed, and the proposed method demonstrated its effectiveness in maximising the production and minimising the total amount of chemical concentrations involved.

**Keywords:** biochemical systems production, constrained optimisation, computational intelligence, cooperative co-evolutionary algorithm, genetic algorithm, Newton method

## 1. Introduction

Computational systems biology is a field of biological study that combines the knowledge of science and engineering. The objective of this field is to model the behaviour of biochemical reactions through a computational approach. Within this field, the structures and complexity of biological processes can be investigated as a system [1]. Therefore, computational systems biology enables the scientist to represent the biological process as a system. This allows the biochemical process in a

living cell to be manipulated as a real factory and gives a way for scientists to improve the cell production (microbial production).

Integrating the knowledge of microbial production with genomic techniques and biotechnology processes creates the ability to manipulate a living cell to act like a real cell factory, thus opening new doors for researchers seeking to improve microbial productions [2]. One example of improving the microbial production is the optimisation of a biochemical systems production. Generally, biochemical systems can be defined as a series of chemical reactions found in a microorganism cell. With the knowledge of microbial production and genomic techniques, the biochemical systems can be represented as a dynamic mathematical model such as the Michaelis-Menten type [3], the stoichiometric approach [4], flux-balance analysis [5], metabolic control analysis [6] and biochemical systems theory (BST) [7]. Among these various choices, this work uses the BST representation to model the biochemical system. An advantage of using the BST is that prior knowledge of the mechanisms for each reaction is not required in order to build equations and the mathematical models can be designed by identifying the reactants and their interconnections [7]. For that reason, a canonical form that uses an ordinary differential equation (ODE) representation is suitable for modelling biochemical systems [1].

The optimisation of the biochemical systems production is a biotechnological process that aims to improve production by fine-tuning the chemical reaction. Besides that, the total amount of chemical concentrations involved also needs to be taken into account [8, 9]. To date, many studies have been carried out to develop methods for the optimisation of the biochemical systems production. Researchers tend to use the computational methods due to the flexibility of the mathematical models allowing to reduce the required costs and time. Popular methods used are the linear programming method (Vera et al. 2010; Xu 2012) and the geometric programming method [10, 11]. These methods depend on the definitions of the decision variables and the equality and inequality constraints, which could cause a convergence problem if the definition process is not performed well [12]. In order to overcome this problem, the present study was carried out using the stochastic method. The stochastic method operates on an evolving set of candidate solutions. In the evolving process, the candidate solutions are modified by the stochastic operator to produce the next generation. Using the stochastic operator, the search direction is determined by a random method, which makes it more efficient and robust [13]. In addition, the stochastic method does not rely on the manipulation of the objective function and constraints or the initialisation of a feasible point [14]. There are many stochastic methods that can be adopted for the optimisation process, among which is the genetic algorithm (GA) that has been widely found to be the most suitable method [15–17]. The GA works by representing the chemical reaction in the biochemical systems as a chromosome. The chromosome is then evolved and modified by a crossover and mutation process, with the intention to improve the solution.

As mentioned above, this chapter uses the BST method to model biochemical systems. Within the BST, two representations are typically used, namely, the S-system and generalised mass action (GMA). This study employs the GMA representation due to its ability to represent the nonlinearity of a biochemical systems and superior performance in optimisation [10]. The GMA uses the power law function, which is an ODE to model the biochemical systems. Applying only the GA for the optimisation of biochemical systems is not sufficient as the GA only fine-tunes the chemical concentrations. Therefore, a method is needed to deal with the biochemical systems. Implementing the Newton method for the biochemical systems is a good choice because the GMA model that represents the biochemical systems can be viewed as a nonlinear equations system [8, 18–22]. It also has been

found that the Newton method is suitable for the nonlinear equations system due to the convergence speed, simplicity and ease of use [23, 24].

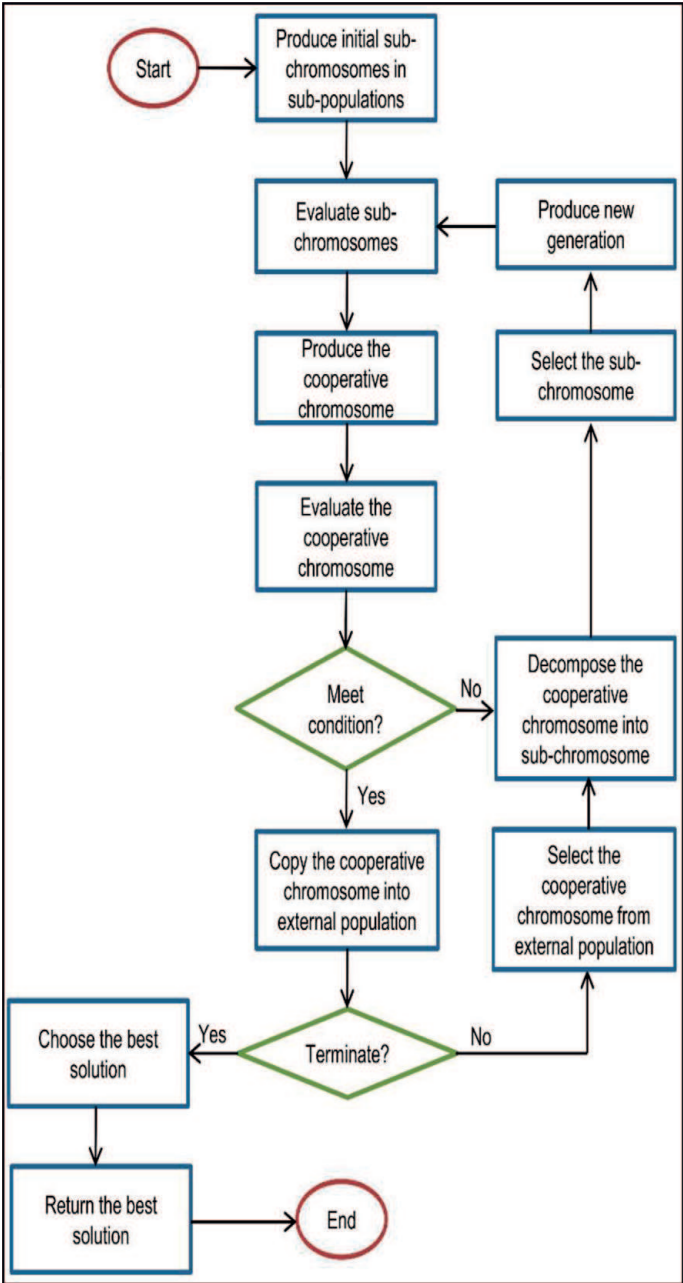
Using the Newton method with the GA in optimising the biochemical systems production is a good choice because the Newton method deals with the biochemical systems, while the GA is used to fine-tune the chemical concentrations by representing the chemical concentrations into a chromosome. However, several problems do occur when dealing with large biochemical systems that contain many chemicals and has complex structures where it makes the representation of the solution become complex and difficult to evaluate. Hence, a method is needed in order to overcome these problems by simplifying the representation of the solution. Using the cooperative co-evolutionary algorithm (CCA) is a good choice because it has the ability to simplify the representation of the candidate solution by decomposing a single chromosome into multiple sub-chromosomes [17, 25, 26].

In this chapter, a hybrid method known as the advanced Newton cooperative genetic algorithm (ANCGA) that combined the Newton optimisation method; the GA and the CCA were presented. This method models biochemical systems as a system of nonlinear equations and applies the Newton method to solve the system. In the optimisation process, the GA and the CCA were used to represent the variables in a nonlinear system in order to search the best solution. The GA was used to maximise the production, while the CCA was used to minimise the total amount of chemical concentrations involved. The ANCGA that proposed in this study is the improvement of the existing method [17]. The reason of proposing the ANCGA is due to the previous algorithm that takes longer time for the optimisation process. Moreover, the performance of the previous work can be improved in terms of maximising the production and minimising the total amount of chemical concentrations involved. In order to do that, this work introduces a concept of external population. The external population was used to store the best solution found in every generation. The reason of using this concept was to avoid the best solution found in every generation from being lost during the reproduction process. The methods used in this study are presented in the following order. Firstly, the proposed method is explained in detail. Case studies of the fermentation pathway in *Saccharomyces cerevisiae* (*S. cerevisiae*) and the *tryptophan* (*trp*) of biosynthesis in *Escherichia coli* (*E. coli*) are then presented. Following that, the results are discussed, and a brief conclusion is made.

## 2. The proposed method

This section describes the proposed ANCGA in detail. The ANCGA is proposed in order to improve the performance of the previous method [17] in terms of computational time. In addition, the ANCGA is hope to improve the performance of the previous method [17] in maximising the production and minimising the total amount of chemical concentrations involved. **Figure 1** shows the flowchart of ANCGA. The ANCGA operates by treating the biochemical systems as a system of nonlinear equations and then uses the Newton method in solving the nonlinear equations system. Then, the GA and CCA were used in the optimisation process. The detailed operation of the ANCGA is described in the following steps:

Step 1—randomly generate the initial  $n$  sub-chromosomes in  $m$  sub-populations and create an empty external population. The number of sub-populations ( $m$ ) must be the same to the number of variables in the nonlinear equations system. The sub-chromosomes represent the variables in the nonlinear equations system. The sub-chromosome is in the binary format.



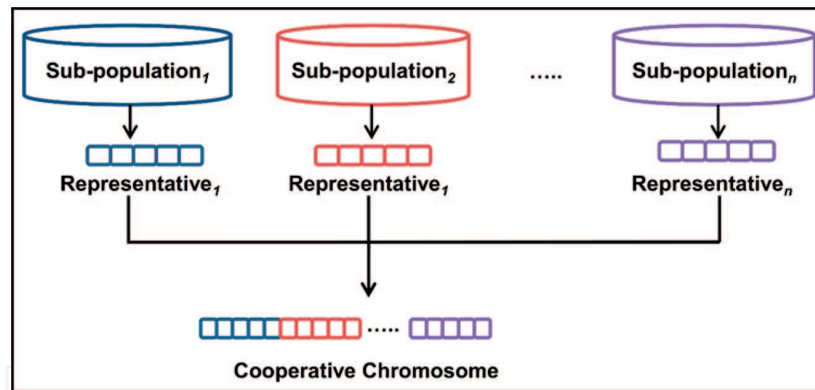
**Figure 1.**  
The flow chart of ANCGA.

Step 2—evaluate the sub-chromosome. The evaluation process starts when a representative from every sub-population is selected to produce a complete solution that is known as a cooperative chromosome. The selection of representatives is based on their fitness value, where the lowest values are selected first. This process is known as the sub-chromosome evaluation. The objective of this process is to minimise the total amount of chemical concentrations involved by letting representatives that have the lowest fitness values from every sub-population to be combined together.

Step 3—produce the cooperative chromosome. The cooperative chromosome is produced after all the selected representatives are combined together. The cooperative chromosome is the complete solution. The formation of the cooperative chromosome is depicted in **Figure 2**.

Step 4—evaluate the cooperative chromosome. In this step, the cooperative chromosome is tested. The evaluation process starts with an encoding of the cooperative chromosome into variables in the nonlinear equations system. Then, the





**Figure 2.**  
 The formation of the cooperative chromosome.

Newton method is used to solve the nonlinear equations system. In the evaluation process, a condition might occur depending on whether or not the cooperative chromosome follows the set of constraints. If the cooperative chromosome follows the constraints, then the procedure goes ahead to Step 8; if not, it goes to Step 5.

Step 5—decompose the cooperative chromosome into sub-chromosomes. After solving the nonlinear equations system using the Newton method, the variables in the nonlinear equations system are decoded back into the cooperative chromosome form. Then, the cooperative chromosome is decomposed into multiple sub-chromosomes. After that, all the sub-chromosomes are sent back to their own sub-populations in order to perform selection and reproduction.

Step 6—select a pair of sub-chromosome for the reproduction process. The selection process is based on their fitness value, where the lowest fitness value is selected first.

Step 7—produce new generations. In this step, the genetic operators of crossover and mutation are applied on the selected sub-chromosomes in order to produce new generations. This process is performed up to the last sub-chromosome. Then, the new generation is processed again, starting from Step 2.

Step 8—copy the cooperative chromosome into the external population. The process is performed by copying selected cooperative chromosome that fulfil the constraints and put the selected cooperative chromosome into external population. This process is intended to keep the best solution in every generation and prevent it from being lost in the reproduction process (Step 7). At this stage, two conditions may occur: either the maximum number of generations is reached or the maximum number of cooperative chromosomes in the external population is achieved. If these two conditions are fulfilled, the procedure jumps to Step 10; otherwise, the procedure continues to the next step. During this process, if the maximum number of cooperative chromosomes in the external population is reached before the maximum number of generations is achieved, the cooperative chromosome that has the lowest fitness value is deleted and replaced by a newly copied cooperative chromosome. However, if the maximum number of generations is reached before the maximum number of cooperative chromosomes in the external population is achieved, the procedure moves to Step 10.

Step 9—select some of the cooperative chromosomes from the external population. This process refers to the elitism of external population concept. The elitism of external population concept works where some (with  $y$  probability) of the cooperative chromosomes from the external population are selected and combined with the current sub-chromosomes. The selection process is based on their fitness value, where the cooperative chromosomes from the external population that have the highest fitness value are selected first. Then, this process goes back to Step 5.

Step 10—choose the best solution. The best solution is chosen among all the cooperative chromosomes in the external population. The selection is based on the fitness values of the cooperative chromosomes, where the cooperative chromosome with the highest fitness value is chosen.

Step 11—return to the best solution. This step decodes the selected cooperative chromosome into its real value (the variable in the nonlinear equations system) and gives the best solution set.

### 3. Case studies

In this section, the effectiveness and efficiency of the ANCGA is demonstrated. The effectiveness of the proposed method refers to the ability of the ANCGA to obtain the best result, while the efficiency refers to the ability of the ANCGA to maintain its performance in producing the best result in several case studies. Two case studies were used, namely, the *S. cerevisiae* pathway and the *E. coli* pathway. In order to test the performance of the ANCGA, a Java program based on two Java libraries, namely, jMetal [27] and JAMA of the version 1.0.3, was developed. The jMetal library can be downloaded from <http://jmetal.sourceforge.net/index.html>, while the JAMA version 1.0.3 can be accessed at <http://math.nist.gov/javanumerics/jama/>.

#### 3.1 Case study 1: the ethanol production in *S. cerevisiae* pathway

In this case study, the ANCGA was used to optimise ethanol production in the *S. cerevisiae* pathway. The GA was used to represent the chemical reactions in the *S. cerevisiae* pathway, which were metabolites and enzymes. Details of the metabolites and enzymes, including the initial steady-state values, are presented in **Table 1**. The pathway was suspended in a cell culture at p. 4.5 and had the following ODE models [28].

Metabolite/enzyme	Symbol	Initial steady-state value
Glc <sub>in</sub>	X <sub>1</sub>	0.0345
G6P	X <sub>2</sub>	1.0110
FDP	X <sub>3</sub>	9.1440
PEP	X <sub>4</sub>	0.0095
ATP	X <sub>5</sub>	1.1278
V <sub>in</sub>	Y <sub>1</sub>	19.70
V <sub>HK</sub>	Y <sub>2</sub>	68.50
V <sub>PFK</sub>	Y <sub>3</sub>	31.70
V <sub>GAPD</sub>	Y <sub>4</sub>	49.90
V <sub>PK</sub>	Y <sub>5</sub>	3440.00
V <sub>Carb</sub>	Y <sub>6</sub>	14.31
V <sub>Gro</sub>	Y <sub>7</sub>	203.00
V <sub>ATPase</sub>	Y <sub>8</sub>	25.10

**Table 1.**  
Details of metabolite and enzymes in case study 1.

$$\begin{aligned}
 \frac{dX_1}{dt} &= V_{in} - V_{HK} \\
 \frac{dX_2}{dt} &= V_{HK} - V_{PFK} - V_{Carb} \\
 \frac{dX_3}{dt} &= V_{PFK} - V_{GAPD} - 0.5V_{Gro} \\
 \frac{dX_4}{dt} &= 2V_{GAPD} - V_{PK} \\
 \frac{dX_5}{dt} &= 2V_{GAPD} + V_{PK} - V_{HK} - V_{Carb} - V_{PFK} - V_{ATPase}
 \end{aligned} \tag{1}$$

Eq. (2) shows the fluxes at the steady-state condition.

$$\begin{aligned}
 V_{in} &= 0.8122X_2^{-0.2344}Y_1 \\
 V_{HK} &= 2.8632X_1^{0.7464}X_5^{0.0243}Y_2 \\
 V_{PFK} &= 0.5232X_2^{0.7318}X_5^{-0.3941}Y_3 \\
 V_{Carb} &= 8.904 \times 10^{-4}X_2^{8.6107}Y_6 \\
 V_{GAPD} &= 7.6092 \times 10^{-2}X_3^{0.6159}X_5^{0.1308}Y_4 \\
 V_{gro} &= 9.272 \times 10^{-2}X_3^{0.05}X_4^{0.533}X_5^{-0.0822}Y_7 \\
 V_{PK} &= 9.471 \times 10^{-2}X_3^{0.05}X_4^{0.533}X_5^{-0.0822}Y_5 \\
 V_{ATPase} &= X_5X_8
 \end{aligned} \tag{2}$$

For the total amount of chemical concentration involved, it can be formulated as follows:

$$\min F_2 = \sum_{j=1}^5 X_j + \sum_{k=1}^6 X_k \tag{3}$$

In the optimisation process, the GMA model was treated as a nonlinear equations system, where all the GMA models were set to be equal to 0. This gave all the ODE models in Eq. (1) the following forms:

$$\begin{aligned}
 V_{in} - V_{HK} &= 0 \\
 V_{HK} - V_{PFK} - V_{Pol} &= 0 \\
 V_{PFK} - V_{GAPD} - 0.5V_{Gol} &= 0 \\
 2V_{GAPD} - V_{PK} &= 0 \\
 2V_{GAPD} + V_{PK} - V_{HK} - V_{Pol} - V_{PFK} - V_{ATPase} &= 0
 \end{aligned} \tag{4}$$

For the metabolite concentration constraint, the constraint was set to 20% from the steady-state value, which was in the range between 0.8 and 1.2 [8, 29]. Thus, the constraint for this case study became as follows:

$$X_k^{0.8} \leq X_k \leq X_k^{1.2} \quad k = 1, 2, 3, 4, 5 \tag{5}$$

Meanwhile, the enzyme concentration constraint was set in the range between 0 and 50 from the steady-state value [8, 29]. The enzyme concentration constraint can be formulated as follows:

$$Y_k^0 \leq Y_k \leq Y_k^{50} \quad k = 1, 2, 3, 4, 5, 8 \tag{6}$$



### 3.2 Case study 2: the tryptophan biosynthesis in *E. coli* pathway

For case study 2, the ANCGA was used to optimise the end product of this pathway, which was *trp* production. The complete description of this pathway was provided by Xiu and colleagues [30]. The GMA models of this pathway are given as follows:

$$\begin{aligned}\frac{dX_1}{dt} &= V_{11} - V_{12} \\ \frac{dX_2}{dt} &= V_{21} - V_{22} \\ \frac{dX_3}{dt} &= V_{31} - V_{32} - V_{33} - V_{34}\end{aligned}\quad (7)$$

where  $X_1$  is the mRNA concentration,  $X_2$  is the enzyme concentration and  $X_3$  is the *trp* concentration. The rates of all reactions in this pathway at steady state are given as follows:

$$\begin{aligned}V_{11} &= 0.6403X_3^{-5.87 \times 10^{-4}}X_5^{-0.8332} \\ V_{12} &= 1.0233X_1X_4^{0.0035}X_{11}^{0.9965} \\ V_{21} &= X_1 \\ V_{22} &= 1.4854X_2X_4^{-0.1349}X_{12}^{0.8651} \\ V_{31} &= 0.5534X_2X_3^{-0.5573}X_6^{0.5573} \\ V_{32} &= X_3X_4 \\ V_{33} &= 0.9942X_3^{7.0426 \times 10^{-4}}X_7 \\ V_{34} &= 0.8925X_3^{3.5 \times 10^{-6}}X_4^{0.9760}X_8X_9^{-0.0240}X_{10}^{-3.5 \times 10^{-6}}\end{aligned}\quad (8)$$

The *trp* production in this case study is given by the reaction  $V_{34}$  [31]. This leads to optimisation that can be formulated as follows:

$$\max F = V_{34}\quad (9)$$

For the total amount of chemical concentrations involved, it can be formulated as follows:

$$\min F_2 = \sum_{k=1}^5 X_k + X_8\quad (10)$$

Similar to case study 1, the GMA model was set to be equal to 0, thus Eq. (8) became as follows:

$$\begin{aligned}V_{11} - V_{12} &= 0 \\ V_{21} - V_{22} &= 0 \\ V_{31} - V_{32} - V_{33} - V_{34} &= 0\end{aligned}\quad (11)$$

In this case study, the GA and CCA only represent several chemical concentrations. This was because not all chemical concentrations were being tuned [1, 10, 11]. The chemical concentrations that tuned were  $X_1$  up to  $X_6$  and  $X_8$ . These chemical concentrations including their initial steady states are summarised in **Table 2**. For the other chemical concentrations which were  $X_7$  and  $X_9$  up to  $X_{13}$ , fixed values were used [1, 10, 11]. Eq. (12) lists the range of these chemicals.

Reaction	Initial steady-state value
$X_1$	0.0345
$X_2$	1.0110
$X_3$	9.1440
$X_4$	0.0095
$X_5$	1.1278
$X_6$	19.70
$X_8$	25.10

**Table 2.**  
*Summary of reaction concentrations in case study 2.*

$$\begin{aligned} X_k^{0.8} \leq X_k \leq X_k^{1.2} \quad k &= 1, 2, 3 \\ 0 \leq X_4 \leq 0.00624 \\ 4 \leq X_5 \leq 10 \\ 500 \leq X_6 \leq 5000 \\ X_7 &= 0.0022X_5 \\ 0 \leq X_8 \leq 1000 \\ X_9 &= 7.5 \\ X_{10} &= 0.005 \\ X_{11} &= 0.9 \\ X_{12} &= 0.02 \\ X_{13} &= 0 \end{aligned} \tag{12}$$

4. Results and discussion

In performing the experiments, many parameter settings were used. The list of all parameter settings used in this study is listed in **Table 3**, whereas **Table 4** presents the parameter settings in producing the best result. The binary coding was used to represent the chemical concentrations. For the Newton method, fixed

Parameter	Rate
Number of sub-populations	Depend on the number of variables in nonlinear equations system
Number of sub-chromosomes in sub-population	[100,110,120,130,140,150]
Number of chromosomes in external population $P$	[100,110,120,130,140,150]
Maximum number of generations	[100,110,120,130,140,150]
Crossover rate	[0.1,0.2,0.3,0.4,0.5]
Mutation rate	[0.1,0.2,0.3,0.4,0.5]
Elitism rate	[0.1,0.2,0.3,0.4,0.5]

**Table 3.**  
*List of all parameter settings used.*

Parameter	Case study 1	Case study 2
Number of sub-populations	11	7
Number of sub-chromosomes in sub-population	150	140
Number of chromosomes in external population $P$	100	100
Maximum number of generations	150	130
Crossover rate	0.3	0.4
Mutation rate	0.1	0.1
Elitism rate	0.2	0.2

**Table 4.**  
*Parameter settings in producing optimum solution.*

Variables	Best solution 1	Average
$X_1$	1.1240	0.9951
$X_2$	1.0322	1.0018
$X_3$	0.9900	1.0053
$X_4$	1.1407	1.1297
$X_5$	1.0001	0.9831
$Y_1$	49.8103	49.9793
$Y_2$	45.3702	45.0767
$Y_3$	45.3452	49.8103
$Y_4$	48.5112	47.4064
$Y_5$	49.4448	49.3426
$Y_8$	49.7563	49.7876
$F_1$	53.0200	52.7499
$F_2$	293.5249	294.5178

**Table 5.**  
*The full result of case study 1.*

parameters were used, namely, 50 for the maximum number of iterations and  $10^{-6}$  for tolerance.

The full results obtained by the ANCGA when applied on *S. cerevisiae* pathway are given in **Table 5**. At the best solution, the ANCGA was able to increase the  $F_1$  (ethanol production) up to 53.02 bigger than its initial steady-state value. For the  $F_2$  (total amount of chemical concentrations involved), the proposed method was able to reduce it to 293.5249. All metabolites and enzymes fulfilled their constraints, with all the metabolites staying in the range of 0.8–1.2, while all the enzymes were in the range of 0–50. The performance of the ANCGA was assessed by comparing the result obtained by ANCGA with other works, and the comparison results are listed in **Table 6**. As shown in the table, the ANCGA produced higher results as compared to other methods. In addition, to verify the results achieved by the ANCGA, an average of 100 independent runs was recorded. The results are summarised in **Table 5**. It shows that the average result for the metabolites and enzymes fulfilled their constraints, whereby they were in their optimum range, thus leading to the conclusion that the ANCGA is able to produce reliable results. It can be said that the ANCGA can produce higher production of ethanol as compared to the methods used in other studies.

Work by	$F_1$	$F_2$
Xu [11]	52.38	297.664
Rodriguez-Acosta et al. [29]	52.31	295.270
Previous method [17]	52.91	294.800
ANCGA	53.02	293.5249

**Table 6.**  
*Comparison with other works for case study 1.*

Variables	Best solution	Average
$X_1$	0.8064	1.0742
$X_2$	0.8046	1.1085
$X_3$	0.8000	0.8000
$X_4$	0.0054	0.0054
$X_5$	4.0116	4.4694
$X_6$	5000	5000
$X_8$	1000	1000
$F_1$	3.9774	3.9616
$F_2$	6006.4280	6007.4575

**Table 7.**  
*The full result of case study 2.*

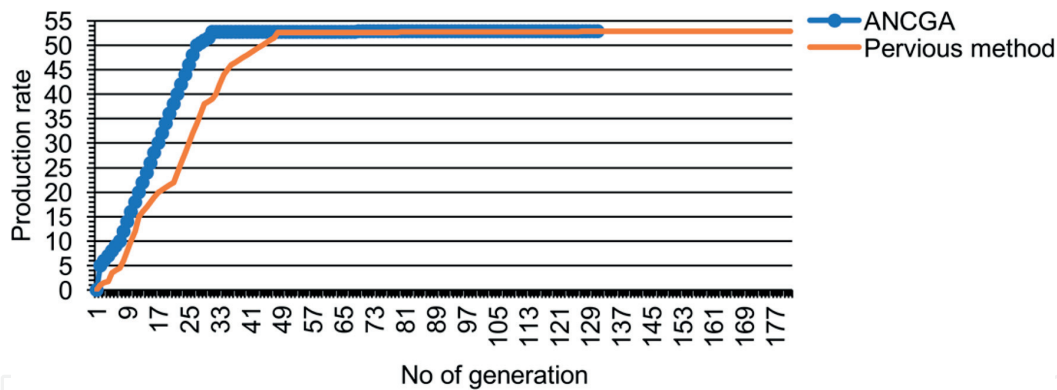
Work by	$F_1$	$F_2$
Marin-Sanguino et al. [10]	3.062	6006.1412
Vera et al. [1]	3.05	6007.1314
Xu [11]	3.946	6007.7814
Previous method [17]	3.9759	6006.5581
ANCGA	3.9774	6006.4280

**Table 8.**  
*Comparison with other works for case study 2.*

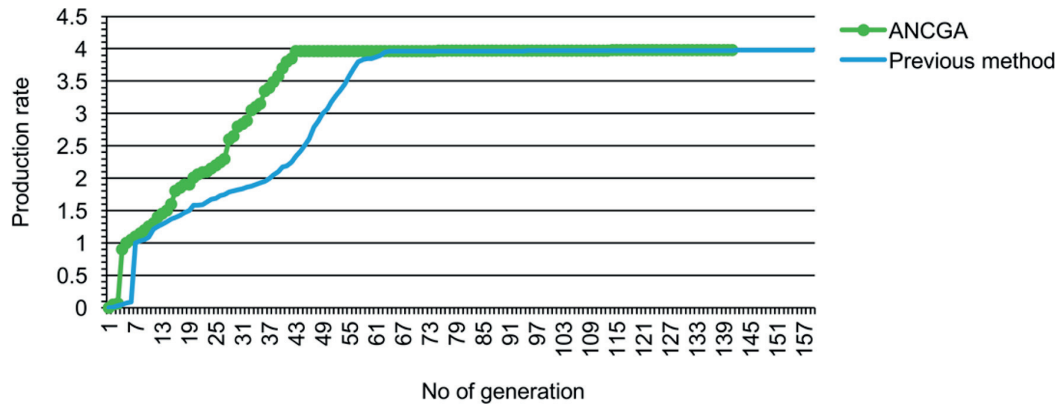
The full results of the *E. coli* pathway are presented in **Table 7**. The ANCGA was able to improve the  $F_1$  (production of *trp*) to 3.9774 from its initial steady state. Meanwhile, the proposed method was able to reduce the  $F_2$  (total amount of chemical concentrations involved) to 6006.4280. All variables representing the chemical reaction followed their constraints and were in the optimum range. To assess the performance of the ANCGA, the results achieved were compared to the results of other methods, with the details of the comparison shown in **Table 8**. As presented in the table, the  $F_1$  of the ANCGA was higher when compared to the methods employed in other works. Similar to the previous case study, 100 experiments were conducted, and the average result was calculated in order to validate the ANCGA results. **Table 7** presents the average result. From the data in **Table 7**, it can be concluded that the ANCGA is reliable in performing the optimisation of this pathway because the average of all the variables follows their constraints. From the observations presented in **Tables 7** and **8**, it can be concluded that the ANCGA is effective in optimising the *trp* production as well as producing reliable results.

The external population concept used by ANCGA can be validated by comparing it with the previous method proposed in [17]. The aim of the external population concept was to reduce the computational time and the number of generations. To learn the effect of the external population concept, several experiments were conducted. To investigate the decrease in the number of generations,  $F_1$  was set to 52.5 for case study 1 and 3.90 for case study 2. After  $F_1$  was achieved, the process was stopped. This helped to investigate which method required more generations in achieving the target production. **Figures 3 and 4** illustrate the comparisons of all case studies. In both figures, the maximum number of the external population was smaller as compared to the maximum number of the previous method in achieving  $F_1$ . This was caused by the concept of external population that was introduced in this study. By using this concept, the best solutions found in the iteration process could be maintained and thus enabled the number of generations to be reduced. In addition, it was found that this concept tended to converge faster than the previous method. This meant that the use of the external population concept allowed faster search of the best solution. In conclusion, the external population concept had an impact in reducing the number of generations and helped in faster convergence as compared to previous methods. To determine the statistical significance between the proposed method and previous methods, the paired t-test and the Wilcoxon signed-rank test were used. The result of the statistical tests showed that all p-values were  $<0.05$ , thus confirming that the proposed method significantly improved the previous method.

Meanwhile, to investigate the decrease in computational time, the maximum number of generations was not set, but  $F_1$  was set to 52.5 for case study 1 and 3.9 for case study 2. After  $F_1$  was achieved, the process was terminated. **Table 9** lists the



**Figure 3.**  
The comparison of results of elitism concept and non-elitism concept for case study 1.



**Figure 4.**  
The comparison of results of elitism concept and non-elitism concept for case study 2.



Method	Case study 1	Case study 2
ANCGA	75.56	38.07
Previous method [17]	80.40	40.45

**Table 9.**  
*The computation times obtained (in second).*

computational time results, and it was found that the ANCGA required less time as compared to the method in [17]. This situation occurred because the high-quality solutions were stored in the external population and then combined with the current solution in the optimisation process. Copying the high-quality solutions into the external population prevented them from being lost (because the optimisation process involving crossover and mutation operation could lose the high-quality solutions). By storing the high-quality solutions into the external population, it would be able to keep the best solution until the optimisation process stopped. To determine the significant improvement of the proposed method against the previous method, the paired t-test and the Wilcoxon signed-rank test were used. The p-value from both tests was  $<0.05$ . From this finding, the proposed method and the previous method were statistically different from each other, and the improvement of the proposed method could be accepted.

## 5. Conclusion

Improving production has become an important issue in the optimisation of biochemical systems. Many factors need to be considered to ensure optimal production. In this work, a hybrid method for constraint optimisation of the biochemical systems production known as the ANCGA was presented. The ANCGA was developed based on a previous method [17], where the ANCGA combined the Newton method, GA and CCA. This study introduced a concept of external population. The aim of this concept was to reduce computational time. In this work, the biochemical system was modelled by a nonlinear equations system. In the optimisation process, the Newton method was employed to deal with a system of nonlinear equations. The GA and CCA were then applied to fine-tune the chemical concentration value in the nonlinear system in order to search for the best solution. During the optimisation process, the high-quality solutions were copied and stored into the external population. The purpose of this process was to avoid the loss of high-quality solutions during the optimisation process. Then, some solutions from the external population were mixed with the next generation of solutions. By doing this, the computational time and number of generations were reduced. In the present study, the proposed method was applied on two case studies, and better results were obtained as compared to the methods presented in other works. In addition, the results were validated, and they demonstrated that the constraints of all the components in the biochemical system were fulfilled. Thus, it can be concluded that the performance of the ANCGA is effective and reliable in producing the best result.

## Acknowledgements

Special appreciation to Universiti Malaysia Pahang for the sponsorship of this study by approving the RDU Grant Vot No. RDU180307. Special thanks to the reviewers and editor who reviewed this manuscript.

IntechOpen

### **Author details**

Mohd Arfian Ismail<sup>1\*</sup>, Vitaliy Mezhuyev<sup>1</sup>, Mohd Saberi Mohamad<sup>2,3</sup>,  
Shahreen Kasim<sup>4</sup> and Ashraf Osman Ibrahim<sup>5,6</sup>

1 Faculty of Computer Systems and Software Engineering, Universiti Malaysia  
Pahang, Gambang, Pahang, Malaysia

2 Institute For Artificial Intelligence and Big Data, Universiti Malaysia Kelantan,  
City Campus, Kota Bharu, Kelantan, Malaysia

3 Faculty of Bioengineering and Technology, Universiti Malaysia Kelantan,  
Jeli Campus, Jeli, Kelantan, Malaysia

4 Soft Computing and Data Mining Centre, Faculty of Computer Science and  
Information Technology, Universiti Tun Hussein Onn, Johor, Malaysia

5 Faculty of Computer Science and Information Technology, Alzaiem Alazhari  
University, Khartoum North, Sudan

6 Arab Open University, Khartoum, Sudan

\*Address all correspondence to: arfian@ump.edu.my

### **IntechOpen**

---

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Vera J, Gonzalez-Alcon C, Marin-Sanguino A, Torres N. Optimization of biochemical systems through mathematical programming: Methods and applications. *Computers & Operations Research*. 2010;**37**(8): 1427-1438
- [2] Sowa SW, Baldea M, Contreras LM. Optimizing metabolite production using periodic oscillations. *PLoS Computational Biology*. 2014;**10**(6): e1003658
- [3] Sakamoto N. Characterization of the transit and transition times for a pathway unit of Michaelis–Menten mechanism. *Biochimica et Biophysica Acta (BBA) - General Subjects*. 2003; **1623**(1):6-12
- [4] Planes FJ, Beasley JE. A critical examination of stoichiometric and path-finding approaches to metabolic pathways. *Briefings in Bioinformatics*. 2008;**9**(5):422-436
- [5] Salleh A, Mohamad M, Deris S, Illias R. Identifying minimal genomes and essential genes in metabolic model using flux balance analysis. In: Selamat A, Nguyen N, Haron H, editors. *Intelligent Information and Database Systems SE - 43*. Vol. 7802. Berlin, Heidelberg: Springer; 2013. pp. 414-423
- [6] Fell D. Metabolic control analysis. In: Alberghina L, Westerhoff HV, editors. *Systems Biology SE - 80*. Vol. 13. Berlin, Heidelberg: Springer; 2005. pp. 69-80
- [7] Voit EO. Biochemical systems theory: A review. *ISRN Biomathematics*. 2013; **2013**:1-15
- [8] Link H, Vera J, Weuster-Botz D, Darias NT, Franco-Lara E. Multi-objective steady state optimization of biochemical reaction networks using a constrained genetic algorithm. *Computers and Chemical Engineering*. 2008;**32**(8):1707-1713
- [9] Xu G. Bi-objective optimization of biochemical systems by linear programming. *Applied Mathematics and Computation*. 2012;**218**(14): 7562-7572
- [10] Marin-Sanguino A, Voit EO, Gonzalez-Alcon C, Torres NV. Optimization of biotechnological systems through geometric programming. *Theoretical Biology and Medical Modelling*. 2007;**4**:38-54
- [11] Xu G. Steady-state optimization of biochemical systems through geometric programming. *European Journal of Operational Research*. 2013;**225**(1): 12-20
- [12] Mariano AP et al. Optimization strategies based on sequential quadratic programming applied for a fermentation process for butanol production. *Applied Biochemistry and Biotechnology*. 2009;**159**(2):366-381
- [13] Balsa-Canto E, Banga JR, Egea JA, Fernandez-Villaverde A, Hijas-Liste GM. Global optimization in systems biology: Stochastic methods and their applications. In: Goryanin II, Goryachev AB, editors. *Advances in Systems Biology*. Vol. 736. New York: Springer; 2012. pp. 409-424
- [14] Mariano AP et al. Genetic algorithms (binary and real codes) for the optimisation of a fermentation process for butanol production. *International Journal of Chemical Reactor Engineering*. 2010;**8**. DOI: 10.2202/1542-6580.2333
- [15] Elsayed SM, Sarker RA, Essam DL. A new genetic algorithm for solving optimization problems. *Engineering*

Applications of Artificial Intelligence.  
2014;27:57-69

[16] Deng H et al. The application of multiobjective genetic algorithm to the parameter optimization of single-well potential stochastic resonance algorithm aimed at simultaneous determination of multiple weak chromatographic peaks. The Scientific World Journal. 2014;2014

[17] Ismail MA, Deris S, Mohamad MS, Abdullah A. A newton cooperative genetic algorithm method for in silico optimization of metabolic pathway production. PLoS One. 2015;10(5): e0126199

[18] Grosan C, Abraham A. A new approach for solving nonlinear equations systems. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans. 2008;38(3): 698-714

[19] Luo Y-Z, Tang G-J, Zhou L-N. Hybrid approach for solving systems of nonlinear equations using chaos optimization and quasi-Newton method. Applied Soft Computing. 2008; 8(2):1068-1073

[20] Babaei M. A general approach to approximate solutions of nonlinear differential equations using particle swarm optimization. Applied Soft Computing. 2013;13(7):3354-3365

[21] Ramos H, Monteiro MTT. A new approach based on the Newton's method to solve systems of nonlinear equations. Journal of Computational and Applied Mathematics. 2017;318:3-13

[22] Ahmad F, Tohidi E, Carrasco JA. A parameterized multi-step Newton method for solving systems of nonlinear equations. Numerical Algorithms. 2016; 71(3):631-653

[23] Liu C-S, Atluri SN. A novel time integration method for solving a large

system of non-linear algebraic equations. Computer Modeling in Engineering and Sciences. 2008;31(2): 71-83

[24] Taheri S, Mammadov M. Solving systems of nonlinear equations using a globally convergent optimization algorithm. Global Journal of Technology & Optimization. 2013;3:132-138

[25] Gu J, Gu M, Cao C, Gu X. A novel competitive co-evolutionary quantum genetic algorithm for stochastic job shop scheduling problem. Computers and Operations Research. 2010;37(5): 927-937

[26] Ismail MA, Asmuni H, Othman MR. The fuzzy cooperative genetic algorithm (FCoGA): The optimisation of a fuzzy model through incorporation of a cooperative coevolutionary method. Journal of Computing. 2011;3(11):81-90

[27] Durillo JJ, Nebro AJ. jMetal: A Java framework for multi-objective optimization. Advances in Engineering Software. 2011;42(10):760-771

[28] Galazzo JL, Bailey JE. Fermentation pathway kinetics and metabolic flux control in suspended and immobilized *Saccharomyces cerevisiae*. Enzyme and Microbial Technology. 1990;12(3): 162-172

[29] Rodriguez-Acosta F, Regalado CM, Torres NV. Non-linear optimization of biotechnological processes by stochastic algorithms: Application to the maximization of the production rate of ethanol, glycerol and carbohydrates by *Saccharomyces cerevisiae*. Journal of Biotechnology. 1999;65(1):15-28

[30] Xiu Z-L, Zeng A-P, Deckwer W-D. Model analysis concerning the effects of growth rate and intracellular tryptophan level on the stability and dynamics of tryptophan biosynthesis in bacteria. Journal of Biotechnology. 1997;58(2): 125-140

[31] Marin-Sanguino A, Torres NV.  
Optimization of tryptophan production  
in bacteria. Design of a strategy for  
genetic manipulation of the tryptophan  
operon for tryptophan flux  
maximization. *Biotechnology Progress*.  
2000;**16**(2):133-145

IntechOpen

IntechOpen