Parameters Estimation of Enzymatic Reaction Model for Biodiesel Synthesis by Using Real Coded Genetic Algorithm with Some Crossover Operations

IOP Conf. Series: Materials Science and Engineering **546** (2019) 052006 doi:10.1088/1757-899X/546/5/052006

Parameters Estimation of Enzymatic Reaction Model for Biodiesel Synthesis by Using Real Coded Genetic Algorithm with Some Crossover Operations

Syaiful Anam[1]
[1]Department of Mathematic, Brawijaya University

Corresponding author: syaiful@ub.ac.id

Abstract. Along with the increase in population and industry in many countries, the fuel oil demand also increases. Petroleum exploration on a large scale will accelerate the depletion of petroleum reserves. One alternative to meet fuel needs is the discovery of biodiesel which is renewable alternative energy. Synthesis biodiesel is carried out through an enzymatic reaction. In the enzymatic reaction model making biodiesel, there are parameters that must be estimated. The estimated parameters of the enzymatic reaction model will determine the success of the reaction. The parameter estimation of the enzymatic reaction model can be done using local optimization or global optimization algorithms, but the local optimization algorithm has a major disadvantage, the optimal value obtained is the local optimal value. Genetic algorithms are global optimization algorithms that are capable of working on high-dimensional problems. The success of genetic algorithms is determined by chromosome models, crossover operations, and mutation operations. The use of improper crossover operations often produces local optimum solutions. There are various types of crossover operation, each of which has weaknesses and advantages. This paper studies the parameters estimation of the enzymatic reaction model for biodiesel synthesis by using genetic algorithms with some crossover operation.

1. Introduction

Recently, the environmental changes are increasingly alarming and threatening the continuity of human civilization. The environmental changes cause some negative effects on the existence of man on earth. The enviroment changes are caused by human factors and natural factors. The air pollution, greenhouse gases production, emission of the carbon dioxide, and other harmful gases are the examples of the enviroment changes by human factors [1]. Air pollution becomes a major problem in the recent years. It has given a serious impact on human health and their environment. The sources of air pollution vary widely, for such as the emission of vehicles and industrial activities [2, 3]. On the human health, the air pollution could cause the onset of diseases such as respiratory infections and inflammations, cardiovascular dysfunctions, and cancer [4-7].

One way to reduce air pollution is to create fuel that produces as little exhaust emissions as possible because exhaust emissions can cause greenhouse effects and air pollution. In addition to environmental problems, human life is now experiencing an energy crisis caused by high fuel needs as a result of population growth, rapid economic and industrial progress in several countries is not balanced by the availability of petroleum as the main source of fuel. Petroleum takes a long time to reform it, so petroleum is a non-renewable energy source. Meanwhile, the high level of dependence of the world

community on petroleum has led to massive exploration which has caused a rapid depletion of petroleum reserves. The discovery of biodiesel is an alternative solution to the problem above because biodiesel is able to reduce exhaust emissions and is renewable alternative energy.

The use of biodiesel can be a solution for Indonesia to reduce its dependence on imported diesel fuel. Biodiesel has also been proven to be environmentally friendly because it does not contain sulfur [8]. In addition, biodiesel is one of a variety of promising alternative renewable energy sources to be developed, because biodiesel feedstock is available abundantly domestically.

Biodiesel synthesis can be carried out through an enzyme reaction called an enzymatic reaction to biodiesel synthesis. Mathematically the enzymatic reaction of biodiesel synthesis can be expressed as a system of ordinary differential equations. The enzymatic reaction model contains parameters that must be estimated, the estimated parameter (parameter estimation) of the enzyme determines the suitability between the model and experimental data. The better the parameters, the more suitable between the model and reality. The problem of parameter estimation of a nonlinear dynamic system is expressed as a problem minimizing functions that measure the suitability of the model with the set of experimental data. This optimization problem can be solved using local optimization algorithms. However, this approach does not usually work for realistic applications, because there is a major disadvantage, namely the dependence of the minimum point solution on the initial value, so that the method becomes very inefficient [9]. Global optimization algorithms have been developed to improve efficiency because this algorithm works well on functions that are not convex, non-smooth and have many local minimums. One of the global optimization algorithms that can be used is the genetic algorithm. Genetic algorithms have advantages in their ease of use and are able to get solutions quickly especially for high dimensional problems. One type of genetic algorithm is real-coded genetic algorithm which has several advantages, namely requiring less storage, working faster, more accurately compared to Binary Genetic Algorithm [10]. The success of genetic algorithms is determined by chromosome models, crossover operations, and mutation operations. The use of improper crossover operations often produces local optimum solutions. The use of improper crossover operations often produces local optimum solutions. There are various types of crossover operation, each of which has weaknesses and advantages.

The paper explains how to estimate the parameters of the enzymatic reaction model of biodiesel synthesis using the real coded genetic algorithm with various types of crossover operation. This papers also analyses the results of parameter estimation of the enzymatic reaction model of biodiesel synthesis using the real-coded genetic algorithm.

2. Research Method

The stages of the research carried out in this study are as follows

1. Literature study
   The first step is that we study the theory of enzymatic reaction models for biodiesel synthesis, parameter estimation and real-coded genetic algorithm
2. Generating data
   The second step is that the research data is generated by using the Runge Kutta method. The Runge Kutta method used is the 4th order Runge Kutta method, because this method has a fairly good error order.
3. Design the real coded genetic algorithm for parameter optimization
   The third step is that we develope the real coded genetic algorithm for parameter estimation of enzymatic reaction models for biodiesel synthesis.
4. Program implementation
   We implement the proposed method by using MATLAB software.
5. Simulation and evalution
   a. Simulation
      This step, we do several simulation for a set of parameters of genetic algorithm. Several cross over operator are applied here.
   b. Analysis of results

The accuration of the methods is evaluated by using Sum Square Error (SSE)

The pseucode of the real coded genetic algorithm for parameter optimization is the following:

Input : Data, genetic algorithm parameter (selection rate, number of population, mutation rate, maximum number of generation, stagnant parameter)
Output : Parameter result from estimization

a) Initialize Population
b) t=1
c) while (t<=number of generation) and (stagnant is False)
    a. Solve the system of ordinary differential equation from each chromosome by using Runge Kutta 4th method
    b. Evaluate Population (calculate the Sum Square Error)
    c. Select Parents by using selection rate and the weighting cost selection)
    d. Do Crossover by using crossover operator
    e. Do Elitism
    f. Do Mutation
    g. Population Replace(Population, Children)
    h. Check the stagnantion
    i. Increase t
   End
d) Get Best Solution from Population
e) Return

3. Results and Discussions

3.1 Structure of the enzymatic reaction model for the formation of biodiesel

Palm-oil methyl ester is one of the biodiesel formed from natural oil. It has been tested on vehicles in Europe, America and the latest in Thailand. This is one of the chemical processes to obtain palm-oil methyl ester which will be used as a test function in the process of parameter estimation. In Figure 1, triglyceride is reacted with methanol to produce diglyceride, diglyceride is reacted with methanol the second to produce monoglyceride and monoglyceride is reacted with the third methanol to produce glycerol. At each level a molecule of methyl ester is produced so that three molecules of esters and one molecule of glycerol are produced as shown in Figure 1. Each reaction is a reversible reaction with a different speed.

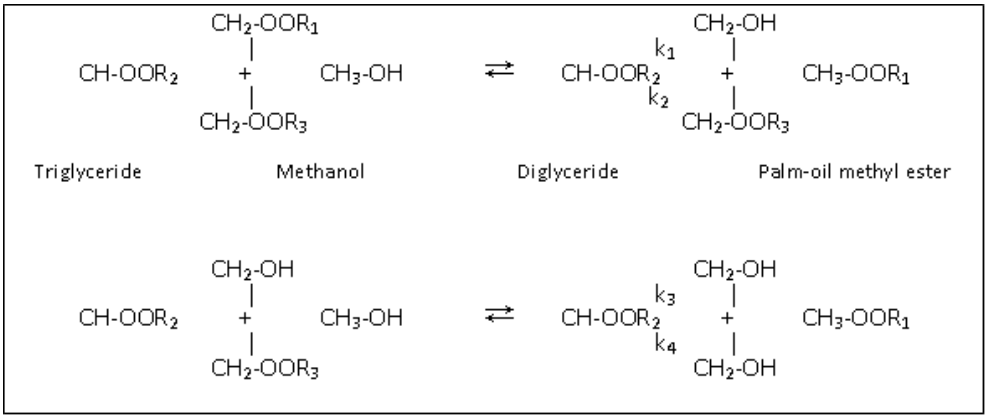

Figure 1. Chemical reactions to obtain palm-oil

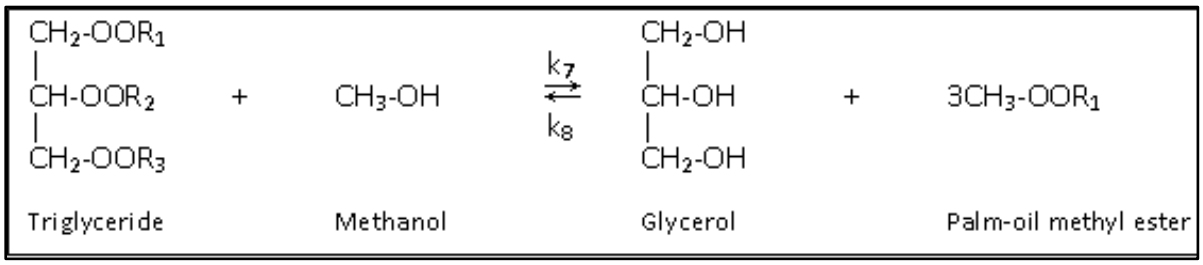

Figure 2. Combining chemical reactions to obtain palm-oil methyl esters

constant (k_n) indicating that the forward reaction and the reverse reaction have different speed constants. The entire chemical process can be modelled in differential equations as follows:

$$\begin{cases} \frac{d[TG]}{dt} = -k_1[TG][A] + k_2[DG][E] \\ \frac{d[DG]}{dt} = k_1[TG][A] - k_2[DG][E] - k_3[DG][A] + k_4[MG][E] \\ \frac{d[MG]}{dt} = k_3[DG][A] - k_4[MG][E] - k_5[MG][A] + k_6[GL][E] \\ \frac{d[GL]}{dt} = k_5[MG][A] - k_6[GL][E] \\ \frac{d[E]}{dt} = k_1[TG][A] - k_2[DG][E] + k_3[DG][A] - k_4[MG][E] \\ \qquad + k_5[MG][A] - k_6[GL][E] \\ \frac{d[A]}{dt} = -\frac{d[E]}{dt} \end{cases} \quad (1)$$

where [TG] is molar concentration of *triglyceride,* [DG]is molar concentration of *diglyceride,* [MG] is molar concenttration of *monoglyceride,* [A] is molar concenttration of *methanol* and [E] is molar concentration of *ester*.

The next step will be discussed about the numerical settlement and parameter estimation of the palm oil transesterfication model with methanol as shown in equation (1), where k_1, k_2, k_3, k_4, k_5 and k_6 are positive parameters [11].

3.2 Simulation data generation

In this study, the data used is data generated by data generation using the Runge Kutta method. Data generation is carried out because of the difficulty of obtaining research data on biodiesel reactions. Data generation is done by providing initial values $k_1 = 0{,}050$, $k_2 = 0{,}11$, $k_3 = 0{,}215$, $k_4 = 1{,}228$, $k_5 = 0{,}242$, $k_6 = 0{,}007$ dan $TG = 1$, $DG = 0$, $A = 6$ in the system of differential equations (1), then the numerical solution using the fourth order Runge-Kutta method with $t = [0;4]$ dan $\Delta t = 0{,}2$.

3.3 Parameter estimation simulation using data generated

Based on the data, the estimation process is carried out using the real-coded genetic algorithm, with parameter values k_1, k_2, k_3, k_4, k_5 and k_6 are located in the interval [0;2]. The parameters of the real-coded genetic algorithm greatly determine the success of the algorithm in solving optimization problems. The method of crossing for the real-coded genetic algorithm used in this study is the blending

method, linear crossover method, one cut point crossover method and heuristic method. Therefore, the determination of existing parameters should not be carried out carelessly because it affects convergence. For the real-coded genetic algorithm, the parameters given are as follows:

- Maximum iteration : 1000
- Population size : 10; 60; 100; 300
- Selection rate : 0.5
- Mutation rate : 0.2
- Stagnantion condition : 100

In this study the experiment was carried out 5 times and the results shown in Table 1 were obtained. This numerical solution for parameter estimation is implemented using MATLAB R2013a and program testing is carried out on computer systems with the following specifications:

Table 1 Comparsion of the SSE of the genetic algorithm with different crossover operators for parameter estimization

Pop Size	One Cut Point Crossover		Blending		Linear Crossover		Heuristic Crossover	
	Means of SSE	Dev. Std of SSE	Means of SSE	Dev. Std of SSE	Means of SSE	Dev. Std of SSE	Means of SSE	Dev. Std of SSE
20	0.0480	0.0621	0.0686	0.0574	0.0446	0.0321	0.0859	0.0702
60	0.0310	0.0163	0.0418	0.0481	0.0133	0.0086	0.0248	0.0097
100	0.0144	0.0161	0.0223	0.0158	0.0115	0.0032	0.0242	0.0088
300	0.0152	0.0031	0.0112	0.0044	0.0127	0.0056	0.0344	0.0097

Table 2. Comparsion of the computational time of the genetic algorithm with different crossover operators for parameter estimization

Pop Size	One Cut Point Crossover		Blending		Linear Crossover		Heuristic Crossover	
	Means of Comp. Time	Dev. Std of Comp. Time	Means of Comp. Time	Dev. Std of Comp. Time	Means of Comp. Time	Dev. Std of Comp. Time	Means of Comp. Time	Dev. Std of Comp. Time
20	30.065	33.9360	27.906	16.6726	29.6283	21.9449	26.6298	18.9001
60	69.122	10.9806	32.955	9.6315	30.6533	5.9744	40.6356	29.0967
100	116.10	69.7087	71.633	10.1114	46.8279	20.3560	113.012	51.0688
300	229.09	67.2544	326.28	151.991	136.047	29.9698	163.687	60.2502

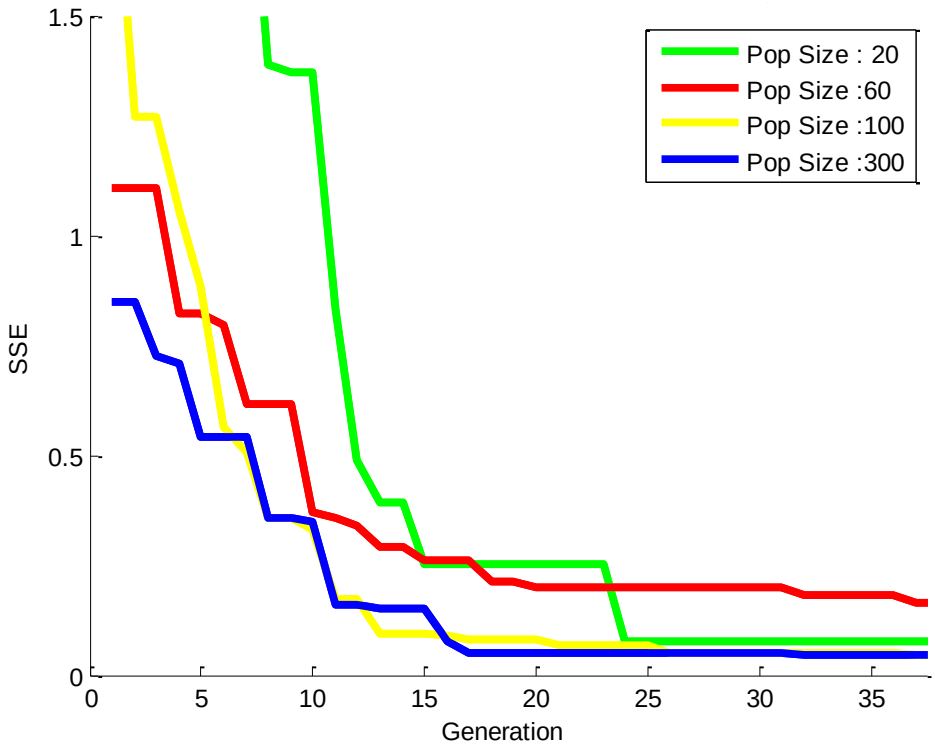

Figure 3. The Relationship between the best SSE and the generation a. Population size of 20, b. Population size of 60, c. Population size of 100, d. Population size of 300

- Intel (R) Core (TM) i5-3317U (1.7 GHz),
- RAM 4GB
- 500 GB hard disk

while the operating system used is the Microsoft Windows 7 Ultimate operating system.

From Table 1, it can be seen that the best estimation results are obtained by using genetic algorithm with linear crossover. The best average SSE value resulted by genetic algorithm with linear crossover operator is 0.0115. The parameters that produce the best fitness are as follows $k_1 = 0.0498$, $k_2 = 0.1009$, $k_3 = 0.1958$, $k_4 = 0.7929$, $k_5 = 0.2917$, and $k_6 = 0.1466$. From Table 2 it can be seen that almost all the larger size of the population will need the longer the computational time. In addition, from Table 1 and 2, it can also be seen that to get the best SEE value does not always take longer than the average computing time. Table 2 also shows that the best means of computational time is almost all obtained by the linear crossover operator.

Figure 3 shows relationship between the best SSE and the generation for population sizes of 20, 60, 100, and 300. It can be seen that the bigger size of population will result the smaller SSE. Based on the estimated parameter values, matching the completion of the system of differential equations to data is presented in Figure 4.

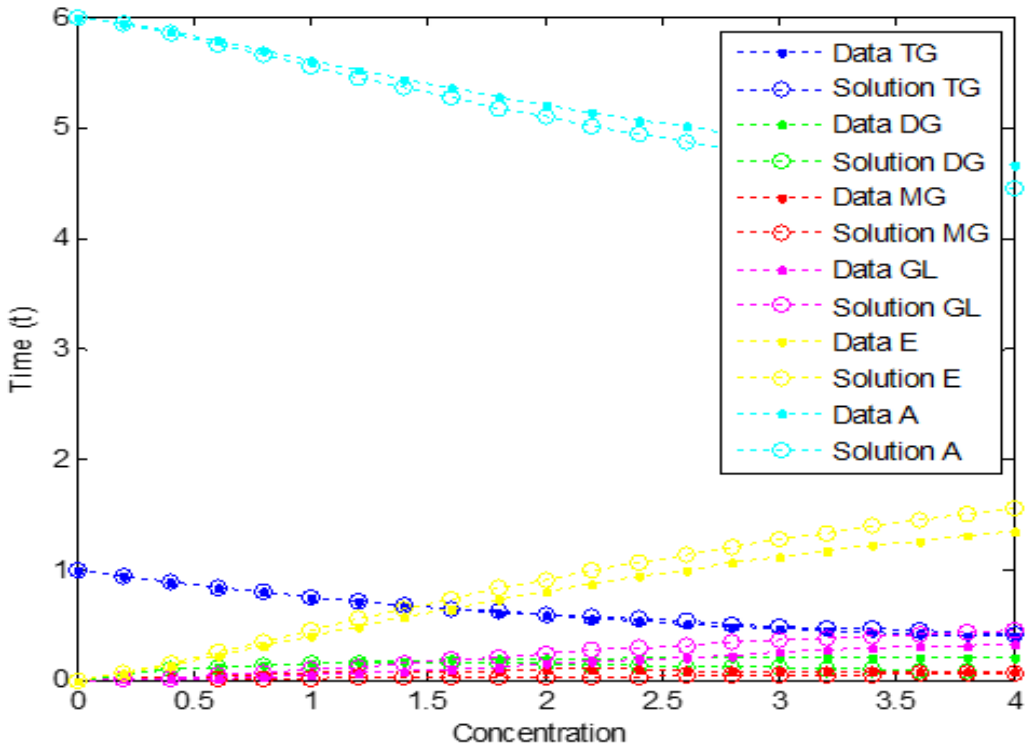

Figure 4. Comparison between data and numerical solutions by using parameter estimated by real-coded genetic algorithm with linear crossover operator.

From the various simulation results obtained, it can be explained that the determination of the real-coded genetic algorithm parameter is very influential on the estimation results. In addition, because all components of the real-coded genetic algorithm are based on random functions (random numbers), so that every time you run a program it will produce different solutions.

From the various simulation results that have been done, the best parameters $k_1 = 0.0498$, $k_2 = 0.1009$, $k_3 = 0.1958$, $k_4 = 0.7929$, $k_5 = 0.2917$, and $k_6 = 0.1466$. This means that there is a match between the model and simulation data or in other words that the reaction model produced is good.

4. Conclusion

Based on the results and discussion that has been done, we can take some conclusions as follows: the real-coded genetic algorithm which is used as a parameter estimation method in the transesterification reaction model is capable of achieving good results, even though the solution is not a global optimum. The linear crossover operator result the smaller SSE compared with the one cut point crossover, heuristic crossover and blending method. Based on the simulation results, the best parameter estimation values are obtained, namely $k_1 = 0.0498$, $k_2 = 0.1009$, $k_3 = 0.1958$, $k_4 = 0.7929$, $k_5 = 0.2917$, and $k_6 = 0.1466$.

References

[1] Nwankwoala H N L 2015 Causes of climate and environmental changes: the need for environmental-friendly education policy in Nigeria Journal of Education and Practice vol 6 no 30 pp 224-234

[2] Robinson D L 2005 Air pollution in Australia: Review of costs, sources and potential solutions. Health Promotion Journal Australia vol 16 pp 213-20.

[3] Habre R, Coull B, Moshier E, Godbold J, Grunin A, Nath A. 2014 Sources of indoor air pollution in New York city residences ofasthmatic children. Journal of Exposure Science and Environmental Epidemiology vol 24 pp 269-278.

[4] Rumana H S, Sharma R C, Beniwal V, Sharma A K 2014 A retrospective approach to assess human health risks associated with growing air pollution in urbanized area of Thar Desert, Western Rajasthan, India. Journal of Environmental Health Science and Engineering vol 12:23.
[5] Yamamoto S S, Phalkey R, Malik A A 2014 A systematic review of air pollution as a risk factor for cardiovascular disease in South Asia: Limited evidence from India and Pakistan. International Journal of Hygiene and Environmental Health vol 217 pp 133-44.
[6] Zhang W, Qian C N, Zeng Y X 2014 Air pollution: a smoking gun for cancer. Chinese Journal Cancer vol 33 pp 173-175.
[7] Brucker N, Charão M F, Moro A M, Ferrari P, Bubols G, Sauer E. 2014 Atherosclerotic process in taxi drivers occupationally exposed to air pollution and co-morbidities. Environmental Research vol 131 pp 31-38.
[8] Anonymous. 2008. Tanggung Jawab Dunia Bisnis Industri. http:// Reuters.blogspot.com/0208/ (February 2008).
[9] Moles C. G, Pedro M and Banga J R 2003 Parameter estimation in biochemical pathways: a comparison of global optimization methods. Genome Research, vol 13 no 11 pp 2467-2474
[10] Haupt, R. L. dan S. E. Haupt. 2004. Practical Genetic Algorithms. John Wiley & Sons, Inc. John Wiley & Sons, Inc.
[11] Allen, M. and G. Prateepchaikul. The modelling of the biodiesel reaction. http://journeytoforever.org/biofuel_library/macromodx.html Department of Mechanical Engineering, Faculty of Engineering, Prince of Songkia University, Thailand.