PAPER • OPEN ACCESS

Grey Wolf Optimizer for Parameter Estimation of Enzymatic Reaction in Biodiesel Synthesis

To cite this article: Syaiful Anam and Indira Kumaralalita 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **546** 052005

Grey Wolf Optimizer for Parameter Estimation of Enzymatic Reaction in Biodiesel Synthesis

Syaiful Anam[1*] and Indira Kumaralalita[1]

[1]Department of Mathematic, Brawijaya University

*Corresponding author: syaiful@ub.ac.id

Abstract. Computational models are used to help us to understand the mechanisms of a complex process in nature. Before building the model, we need to know the characteristic of samples which are described in the form of measure named parameters. Usually, the value of parameters is unknown and we need to investigate those value to know the compatibility between the artificial model and real circumstances. Many optimization methods have been introduced to estimate those parameters, but some of them meet the difficulties caused by the nonlinear type of function model. Many objective functions of the estimation parameter are multimodal, high dimensional, and have many local optima, so the estimation process using traditional optimization method is not suggested. In this article, we use Grey Wolf Optimizer (GWO), as one of the metaheuristic artificial intelligence algorithm, which is inspired by leadership hierarchy and hunting behavior of a pack of wolves. GWO is applied to estimate the parameters in a model of enzymatic reaction in biodiesel synthesis. Biodiesel is renewable fuel that can solve the energy crisis and pollution. While the process of biodiesel synthesis occurs, some enzyme in the biodiesel substances react to each other and it can be modeled into ODEs (Ordinary Differential Equations) system. The kinetic parameters inside them are needed to be estimated. After the parameter are estimated, the fourth-order Runge-Kutta method is used to solve the system. The result is evaluated by analyzing the objective function which minimizes the Sum of Squared Errors (SSE). The small value of SSE and the narrow range of both parameter of model and estimation shows that GWO is effective to be the proposed method for parameter estimation and model selection problems.

1. Introduction

The world is damaged by various things and with no exception by the effect of developing in the industry. In the last 10 years, National Surveys on Energy and the Environment (NSEE) found that the number of Americans trust the existence of solid evidence of global warming becomes higher. They stated that humans are being the most important role in the occurrence of global warming on this earth [1]. Many high-technologies are created by people but the creator did not pay attention to specific things such as the safest method of use and the installation of waste. When the engine is running, exhaust emissions emerge, which is the remaining fuel from the engine. Because of that, the air is covered with these emissions and make pollution. Besides the dangerous pollution, the usage of fuel is not renewable and not managed properly. It made the energy become rare then trigger the energy crisis, especially in Indonesia..

Some actions should be arranged for reducing those severe damages. One of the ways is to look for other resources that can be updated but are kept secure for the earth. Those problems are solved by the discovery of biodiesel which being used starting from 1984. Biodiesel is a fuel derived from renewable vegetable oils or animal fats. Those two substances contain triacylglycerols (TAG) and pass some chemistry process, then, in the end, it converted to be biodiesel [2]. In Indonesia, those substance existences are abundantly so we can renew this fuel.

The kinetics of the enzyme reactions in making biodiesel can be represented in the Ordinary Differential Equation (ODE) system. There are some parameters in that system which are needed to be estimated. The parameter estimation process will be compared by the real value which is given. There are a lot of optimization methods that can estimate the parameter, but some of the methods are failed to reach the optimum global solution because of the nonlinear and non-continuous functions which are consisted in the model. Metaheuristic optimization is developed to repair the previous method. One of the method, named Grey Wolf Optimizer (GWO), inspired by the social and hunting behavior of Grey Wolf in real nature. The advantage of GWO compared to other optimization algorithms, such as Particle Swarm Optimization (PSO) and Gravitational Search Algorithm (GSA), is that GWO is able to solve nonlinear function optimization problems by getting the minimum value with shorter computation time than other algorithms [3].

The goal of the optimization process in this article is to find the optimum parameters of kinetic rate constant value from the transesterification process and minimizes the objective function or the Sum of Squared Error (SSE). The minimum of fitness value shows that the model is a good representation to illustrate the enzymatic reaction in biodiesel synthesis.

2. Data

The product of biodiesel is resulted by some combination process named Transesterification. Diglyceride is obtained by reacting triglycerides and methanol, and Diglyceride reacted by another methanol will produce Monoglyceride. The substances, Monoglyceride, and Methanol are reacted into one composition then form Glycerol. All of those each process are produced a molecule of the fatty acids methyl ester and they are added up with the Glycerol to be the product that we know as the palm-oil methyl ester. Those reactions are reversible with the different rate constant. It means that the forward and reverse reaction go on at a different rate. All the processes can be expressed with a second-order of the differential equation system in Equation (1) [4].

$$\begin{cases} \dfrac{d(TG)}{dt} = -k_1(TG)(A) + k_2(DG)(E), \\ \dfrac{d(DG)}{dt} = k_1(TG)(A) - k_2(DG)(E) - k_3(DG)(A) + k_4(MG)(E), \\ \dfrac{d(MG)}{dt} = k_3(DG)(A) - k_4(MG)(E) - k_5(MG)(A) + k_6(GL)(E), \\ \dfrac{d(GL)}{dt} = k_5(MG)(A) - k_6(GL)(E), \\ \dfrac{d(E)}{dt} = k_1(TG)(A) - k_2(DG)(E) + k_3(DG)(A) - k_4(MG)(E) + k_5(MG)(A) - k_6(GL)(E), \\ \dfrac{d(A)}{dt} = -\dfrac{d(E)}{dt}, \end{cases} \tag{1}$$

The system in Equation (1) represents the molar concentrations change of each substance, such as TG is a triglyceride molar concentrate, DG is a diglyceride molar concentrate, MG is a monoglyceride molar concentrate, A is a methanol molar concentrate, and E is an ester molar concentrate, in t interval. The molar concentration of each substance can be seen in every change of time by firstly know the value of every reaction rate constant (k_n).

The real experiments have been done by [5] and be used to check the performance of GWO in the parameter estimation k. The reaction rates constant (k_n) will be given as the initial values from the research with certain activation energy values, such as $k_1 = 0.050, k_2 = 0.110, k_3 = 0.215, k_4 = 1.228, k_5 = 0.242$, and $k_6 = 0.007$.

3. Methodology

3.1. A fourth-order Runge-Kutta Method

Many real difficult problems in this world are described using the differential equation system. According to [6], the general form that has been known is given below,

$$\frac{dx}{dt} = f(t, x, y),$$
$$\frac{dy}{dt} = g(t, x, y),$$

due to the initial conditions, $x(t_0) = x_0, y(t_0) = y_0$ on the interval $t_0 \leq t \leq t_n$. The variable t is independent, while x and y are dependent. The parameters t_0, x_0, and y_0 are initial values of t, x, and y, and t_n is the ending value of t. Usually, t has the step size which is represented by h and it is discretized to be t such that $t_0, t_1 = t_0 + h, t_2 = t_0 + 2h, \dots, t_n = t_0 + nh$.

To solve this system problem, many algorithms are developed. One of them is Runge-Kutta method to approach the solution. This method is widely used because of its fast and high rate of convergence. The fourth order Runge-Kutta error has the error order $O(h^2)$. Its formula is given by,

$$x_{i+1} = x_i + \frac{1}{6}(f_1 + 2f_2 + 2f_3 + f_4),$$
$$y_{i+1} = y_i + \frac{1}{6}(g_1 + 2g_2 + 2g_3 + g_4),$$

where

$$f_1 = hf(t_i, x_i, y_i),$$
$$f_2 = hf\left(t_i + \frac{h}{2}, x_i + \frac{f_1}{2}, y_i + \frac{g_1}{2}\right),$$
$$f_3 = hf\left(t_i + \frac{h}{2}, x_i + \frac{f_2}{2}, y_i + \frac{g_2}{2}\right),$$
$$f_4 = hf(t_i + h, x_i + f_3, y_i + g_3),$$
$$g_1 = hg(t_i, x_i, y_i),$$
$$g_2 = hg\left(t_i + \frac{h}{2}, x_i + \frac{f_1}{2}, y_i + \frac{g_1}{2}\right),$$
$$g_3 = hg\left(t_i + \frac{h}{2}, x_i + \frac{f_2}{2}, y_i + \frac{g_2}{2}\right),$$
$$g_4 = hg(t_i + h, x_i + f_3, y_i + g_3).$$

3.2. Grey Wolf Optimizer

Grey Wolf (Canida lupus) is one species carnivore that takes the highest position in the food chain and lives gathering in a group. In their social life, grey wolves are divide into four hierarchies clusters that show the level of the position. The highest level called alpha, the second one is called beta, the third one is called delta, and the lowest one is omega. All of each position has its own task in real nature. Alpha wolf becomes the leader that arrange all the activity in a group, especially for hunting. Grey wolf also has a unique way to hunt their prey. It is because of the big size of the prey that makes the wolves have the new hunting step. When the wolf sees its prey, the alpha wolf will lead the group to catch up the prey. After that, they do the encircling behavior, not persecute their prey directly. They will approach it in the safe distance while evaluating the prey [7].

From that illustration in the real jungle, the general form of wolf's behavior in encircling the prey can be modeled and simulated with these mathematical functions, that expressed in Equation (2).

IOP Conf. Series: Materials Science and Engineering **546** (2019) 052005 doi:10.1088/1757-899X/546/5/052005

$$\begin{aligned} \vec{D}(t) &= \vec{C}(t) \circ \overrightarrow{X_p}(t) - \vec{X}(t), \\ \vec{X}(t+1) &= \overrightarrow{X_p}(t) - \vec{A}(t) \circ \vec{D}(t), \end{aligned} \quad (2)$$

$\vec{X}$ is a wolf vector position, $\overrightarrow{X_p}$ is a vector position of the prey, and $\vec{D}$ is a vector that expresses the interval distance between wolves and its prey which contains positive value. The t variable insides interval $\vec{D}$ indicates the iteration that recurs inside the encircling process, while $\vec{A}$ and $\vec{C}$ are the coefficients that can be counted, as shown in these equations (3),

$$\begin{aligned} \vec{A}(t) &= 2\vec{a}(t) \circ \overrightarrow{r_1} - \vec{a}(t), \\ \vec{a}(t) &= \left(2 - \frac{2t}{t_{max}}, 2 - \frac{2t}{t_{max}}\right), \\ \vec{C}(t) &= 2 \circ \overrightarrow{r_2}. \end{aligned} \quad (3)$$

The t variable in Equation (3) is repeated until t_{max} that the value is given by some tests. The value of $\vec{a}$ is linearly decreased from two to zero as long as the t value is iterated. The decreased value $\vec{a}$ interprets the wolf motion while approaching their prey and also influences the value of $\vec{A}$. Suppose that $\vec{A} = (m_i)$ for $i = 1, \ldots, n$, with n is the dimension of function. When doing iterations, the value of $|\ m_i\ | \geq 1$ represents a wolf agent that disperses for prey or in an algorithm aimed at finding a more appropriate solution, whereas if it shows $|\ m_i\ | < 1$ then the exploitation process is done. Exploration and exploitation are the most important things that influence the optimization. Exploration is the process of finding the optimum solution in all spaces in the domain, while searching for the exploitation process focuses on the solution area, where the optimum point is located.

Variables $\overrightarrow{r_1}$ and $\overrightarrow{r_2}$ are randomly valued vectors which are uniformly distributed in a range of values [0,1]. The $\vec{C}$ parameter is made of a random value to expand the range in the exploration process and stay away from the possibility of a final solution falling at the local optimum point. The positions of the alpha, beta and delta wolves, are considered the three best solutions and all the positions of the wolves will be updated following the three best solutions. The equation models of wolf position update is written as follows,

$$\begin{aligned} \overrightarrow{D_\alpha}(t) &= \overrightarrow{C_1}(t) \circ \overrightarrow{X_\alpha}(t) - \vec{X}(t), \\ \overrightarrow{D_\beta}(t) &= \overrightarrow{C_2}(t) \circ \overrightarrow{X_\beta}(t) - \vec{X}(t), \\ \overrightarrow{D_\delta}(t) &= \overrightarrow{C_3}(t) \circ \overrightarrow{X_\delta}(t) - \vec{X}(t), \end{aligned} \quad (4)$$

$$\begin{aligned} \overrightarrow{X_1}(t) &= \overrightarrow{X_\alpha}(t) - \overrightarrow{A_1}(t) \circ \overrightarrow{D_\alpha}(t), \\ \overrightarrow{X_2}(t) &= \overrightarrow{X_\beta}(t) - \overrightarrow{A_2}(t) \circ \overrightarrow{D_\beta}(t), \\ \overrightarrow{X_3}(t) &= \overrightarrow{X_\delta}(t) - \overrightarrow{A_3}(t) \circ \overrightarrow{D_\delta}(t), \end{aligned} \quad (5)$$

$$\vec{X}(t+1) = \frac{\overrightarrow{X_1}(t) + \overrightarrow{X_2}(t) + \overrightarrow{X_3}(t)}{3}. \quad (6)$$

Vectors $\overrightarrow{D_\alpha}$, $\overrightarrow{D_\beta}$, and $\overrightarrow{D_\delta}$ in Equation (4) show the estimation of the distance between the position of wolf $\vec{X}$ and the best position belonging to the wolf alpha $(\overrightarrow{X_\alpha})$, beta $(\overrightarrow{X_\beta})$, and delta $(\overrightarrow{X_\delta})$ in Equation (5). Parameters $\overrightarrow{A_1}, \overrightarrow{A_2}, \overrightarrow{A_3}, \overrightarrow{C_1}, \overrightarrow{C_2}$, and $\overrightarrow{C_3}$ are calculated using Equation (3). All agents will be updated in position with Equation (6). The position and fitness of each agent will be stored and used for the calculations in the next iteration, but only the solution of the alpha wolf is considered the best solution and printed at the end of the iteration [7].

3.3. Estimation Parameter Process

In this experiment, the data are generated first using a fourth-order Runge-Kutta method with the value of time interval $t = [1,10]$ and $\Delta t = 0.01$. Then, the differential equation system will be solved for the second time but with the parameters that have been estimated using Grey Wolf Optimizer. The random generated parameters are through some process until 1000 maximum iterations in the interval [0,5] or the parameters that fit the real data are found. It uses 150 wolf as the parameters of search agents.

Grey Wolf Optimizer algorithm for estimation parameter is written such that,

Input: max iteration, number of agents, upper limit, lower limit, dimensions.

1. Generate the kinetic rate (k) randomly.
2. Determine the maximum iteration and error.
3. While (loop <maxiter && alpha_score >error)
   a. For i = 1: 6
      1) Set the domain boundary to not to exceed the limit.
      2) Complete Runge-Kutta with k_i that has been generated.
      3) Calculate the Sum Square Error (SSE).
   b. Sort SSE from the smallest.
   c. The smallest fitness value becomes the alpha score, the second smallest fitness becomes the beta score, and the third smallest becomes the delta score. The other values represent omega wolf, who will be updated in position depending on the three wolves fitness that had been determined before.
   d. Update k values by doing calculations on Equation (4), (5), and (6).
   e. loop=loop+1.
   End while.
4. Reach the termination criteria.

Output: alpha_score as fitness, alpha_pos as the value of six kinetic parameters value.

4. Result and Discussion

The numerical parameter estimation in this experience is implemented using MATLAB R2013 and the computer system with the specification processor Intel(R) Core(TM) i3-5005U, CPU 2.0GHz, RAM 4GB, and hard disk 500GB. The operation system that is used for evaluating the algorithm is Microsoft Windows 10 Pro 1803 version.

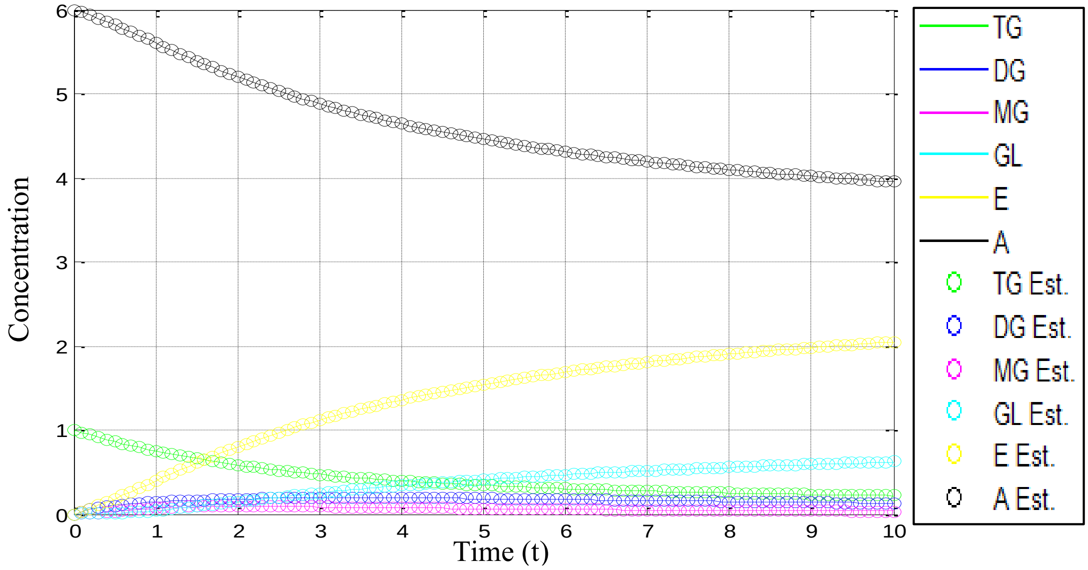

Figure 1. Comparison Between Runge-Kutta Result from Parameter Estimation Process and Real Experiment Data.

IOP Conf. Series: Materials Science and Engineering **546** (2019) 052005 doi:10.1088/1757-899X/546/5/052005

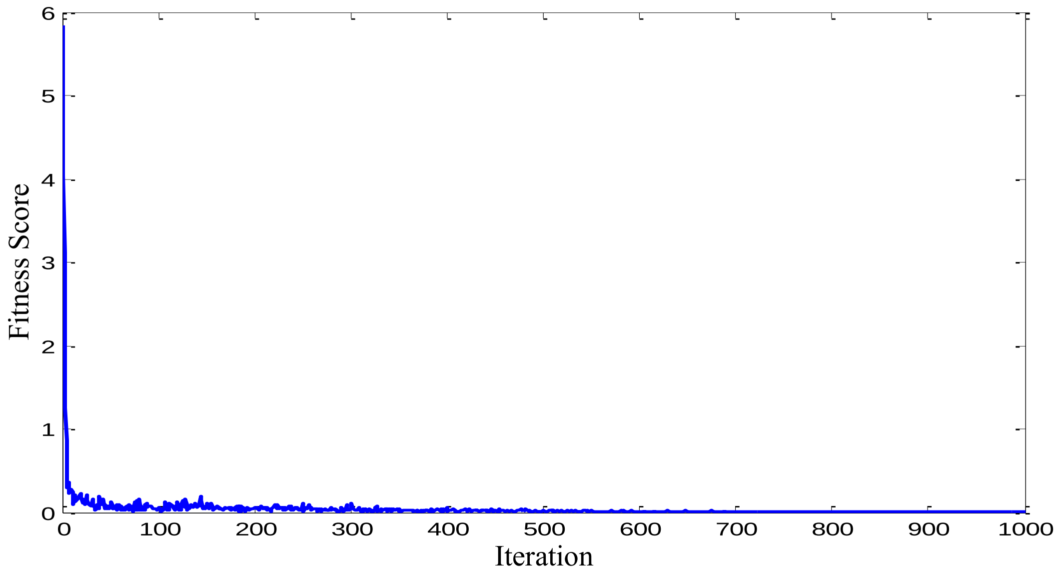

Figure 2. Fitness result from SSE value in 1000 iteration

As the simulation which has been done, the estimation value result of kinetic parameters are $k_1 = 0.0501, k_2 = 0.1113, k_3 = 0.2321, k_4 = 1.4150, k_5 = 0.2291,$ and $k_6 = 0.0000041231$ with the Sum of Squared Error (SSE) value is 0.0023 that show the similarity between real data and the model. The result of estimation parameter process can be seen in Figure 1. From Figure 2, we also know that Grey Wolf Optimizer can have convergent result decreasing along as the iteration progress.

5. Conclusion

The result of the estimated parameter and from the real data are coincided that shown that estimation parameter process is successfully done. It means that there is a compatibility between the data with some parameters get from the experiment and the data which get by doing parameters estimation using Grey Wolf Optimizer with 150 agents and 1000 maximum iteration. According to the result, we can conclude that the estimation parameter process using Grey Wolf Optimizer can reach the good result represents by the minimum fitness (approach the zero). It means that the Grey Wolf Optimizer is suggested to be used in estimation parameter problems because of its performance to find the global optimum solution.

References

[1] Borick, C., Rabe, B.G., Fitzpatrick, N.B., dan Mills, S.B. 2018. As Americans Experienced the Warmest May on Record Their Acceptance of Global Warming Reaches a New High. *Issues in Energy and Environmental Policy (IEEP) Number 37*.

[2] Moser, B.R. 2009. Biodiesel production, properties, and feedstocks. *In Vitro Cell.Dev.Biol.-Plant.* **45** pp.229-266.

[3] Kumaralalita, I. 2019. *Optimasi Fungsi Nonlinear dengan Grey Wolf Optimizer.* Undergraduate Thesis, Brawijaya University. Malang, Indonesia.

[4] Allen, M, and Prateepchaikul, G. 2006. *The modeling of the biodiesel reaction.* Thailand: Department of Mehanical Engineering Faculty of Engineering Prince of Songkla University.

[5] Noureddini, H., and Zhu, D. 1997. Kinetics of Transesterification of Soybean Oil. *Journal of the American Oil Chemists' Society.* **74** pp.1457-1463.

[6] Tay, K.G., Cheong, T.H., Lee, M.F., Kek, S.L., and Kahar, R.A. 2014. A fourth-order Runge-Kutta (RK4) Spreadsheet Calculator For Solving A System of Two First-Order Ordinary Differential Equations Using Visual Basic (VBA) Programming. *Spreadsheets in Education (eJSiE).* 8(1).

[7] Mirjalili, S., Mirjalili, S.M., dan Lewis, A. 2014. Grey Wolf Optimizer. *Advances in Engineering Software.* **69** pp.46-61.