Parameter Estimation for a Model of Ionizing Radiation Effects on Targeted Cells using Genetic Algorithm and Pattern Search Method

Hamizah Rashid*, Fuada Mohd Siam, Normah Maan and Wan Nordiana W Abd Rahman

1, 2Department of Mathematical Sciences, Faculty of Science
Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia
3Medical Radiation Programme, School of Health Sciences
Universiti Sains Malaysia, 16150, Kubang Kerian, Kelantan.
*Corresponding author: hamizahrashid90@gmail.com.my

Abstract A mechanistic model has been used to explain the effect of radiation. The model consists of parameters which represent the biological process following ionizing radiation. The parameters in the model are estimated using local and global optimization algorithms. The aim of this study is to compare the efficiency between local and global optimization method, which is Pattern Search and Genetic Algorithm respectively. Experimental data from the cell survival of irradiated HeLa cell line is used to find the minimum value of the sum of squared error (SSE) between experimental data and simulation data from the model. The performance of both methods are compared based on the computational time and the value of the objective function, SSE. The optimization process is carried out by using the built-in function in MATLAB software. The parameter estimation results show that genetic algorithm is more superior than pattern search for this problem.

Keywords Parameter estimation; Genetic Algorithm; Pattern Search; Double-Strand Break (DSB); Linear-Quadratic formulation.

Mathematics Subject Classification 92B05, 62F10, 65L99.

1 Introduction
Radiation helps a lot in improving human health. In cancer treatment, radiation therapy works well along with surgery and chemotherapy. The first cancer treatment using radiation was recorded in 1896, six months after X-rays was discovered [1]. Since then, the radiotherapy has been evolved in line with time and technology.

However, exposure of cells to radiation can alter the deoxyribonucleic acid (DNA) structures such as breakage on the DNA strand or base missing in the DNA backbone. It will subsequently
cause damage to living cells and causing death. An individual cell can suffer a million alterations of DNA in a day due to DNA strand breakage [2].

The most toxic lesions produced by radiation is double-strand breaks (DSBs). This lesion formed at both sides of the DNA double helix and must be repaired to preserve the chromosomal integrity [3].

In clinical practice, Linear Quadratic (LQ) formulation has been used to determine the survival of targeted cells affected by radiation [4]. The relation of cell survival fraction with respect to radiation dose is given as

\[
S(D) = e^{-\left(\alpha D - \beta D^2\right)}
\]

where \(S(D)\) is the survival fraction of cells with respect to dose, \(D\), \(\alpha\) is a constant that describes the initial slope of the cell survival curve and \(\beta\) is a constant that describes the quadratic component of cell killing. Equation 1 can be written in logarithmic form as follows:

\[
\ln S = -\alpha D - \beta D^2
\]

The value of \(\alpha/\beta\) ratio is fundamental to the LQ formulation. From a biological point of view, the ratio of \(\alpha/\beta\) is related to the ability of the cell to repair. A tissue with low \(\alpha/\beta\) ratio has a higher capacity for self-repair than one with a high ratio [5,6].

2 Mathematical Modelling

In radiobiology, many models have been established and well studied to explain the radiation effect [7,8]. For example, Wentz et al. [9] presented a mathematical model that describes the effect of acute radiation exposure on thrombopoiesis in the human body.

In 2016, Siam et al. [10] developed a mechanistic model of ionizing radiation damage to DNA. The model considered cell population with a different number of double-strand break (DSB) and the misrepair cell due to ionizing radiation. The model then fitted to the LQ formulation to estimate the ratio of \(\alpha/\beta\). Next, they continue their work with parameter estimation using Nelder-Mead (NM) simplex and Genetic Algorithm (GA) method [11]. From their results, Nelder-Mead simplex provides the lowest value of the objective function compared to Genetic Algorithm.

2.1 Curve Fitting and Parameter Estimation

In general, most mechanistic models in system biology are dynamical system. Dynamical systems usually depend on parameters provide physical meaning, which can explain the behavior of the system, such as the death rate of a cell population, the repair rate of the double-strand breaks (DSBs) damage and many more.

Model parameters involve in the biology especially in radiobiology are essential in radiotherapy planning and radiotherapy safety. For example, in the early step in preparing a patient for treated with Ionizing Radiation (IR), the radiation oncologist has to determine the specific parameter, so that the patient received an effective therapy.

One of the challenges in computational modelling of biological system is to determine the value of model parameters. Often, the parameter can be evaluated experimentally. However,
there are some parameters which cannot be measured from the experimental directly. These parameters can be estimated through the curve fitting process.

There are two common approaches in estimating parameter: maximum likelihood and least square. In this research, we only consider the least square method since the probability distribution of the experimental data is unknown. Least square criterion assumed that the best fit line of the experiments data is the line which has the smallest value of the least square error (or sum of the squares of the deviation) from the given data.

In this research, optimization is implemented to find the best fit between experimental data and model data. Objective function is used to represent the sum of squares error (SSE). Optimization is employed in order to find the minimum value of the objective function.

The objective function is the sum of the square error (SSE) between experimental data and simulation data which is given below

\[
SSE = \sum_{i=1}^{n} [y - f(x, \theta)]^2
\]

where \( y \) is the experimental data and \( f(x, \theta) \) is the surviving fraction from the mathematical model that contain parameters. In order to minimize the SSE, two types of optimization methods which are in the category of local and global optimizer are employed.

Local search optimization is a direct solution that moves to its neighborhood iteratively. A new solution is found at each step they are moving. The way of choosing the new solution is to pick up the best one in the neighborhood and replace the current solution and the search continues. For example, \( f(x) \) will be a new solution if the value of \( f(x) > f(x_{n+1}) \) where \( f(x) \) is the value of objective function. The search will quit when the current solution does not have any improvement. The example of local search optimization is Nelder-Mead simplex [12], and Pattern Search [13].

In contrast, global search optimization use gradient of the objective function to determine the direction. It searches solution among all the possible solution which is not restricted to the neighbourhood region only. The example of well-known global optimization is Simulated Annealing [14], Genetic Algorithm [15] and Particle Swarm Optimization [16].

Parameter estimation is performed either using one or two optimization methods simultaneously. For example, Costa et al. focused on global optimization method, Simulated Annealing to estimate parameters value in thermodynamic models [17] as well as Özsoy et al. focused on Particle Swarm Optimization [18].

On the other hands, some researchers perform two different optimization methods to compare and determine which method best fit to experimental data [19,20]. For example, research by Ruf et al. [21] applied both local and global optimization methods to estimate parameter’s value. In their work, local optimization is able to find the best solution between experimental data and their model compared to global approach [21]. Then in 2016, Siam et al. [10] compare the performance between local and global optimization method, Nelder-Mead Simplex and Genetic algorithm respectively. From their result, local optimization method, Nelder-Mead simplex provides the smallest objective function value, sum square error, SSE which indicates the best solution to the problem [11].

A mathematical model of radiation effects to the cell was developed by Siam et al. [10] is employed as a referenced model in this study. As previously mentioned, they compared
the performance of optimization method between Genetic Algorithm (GA) and Nelder-Mead Simplex method [11]. However, in this study, the parameters in the model will be estimated by using local and global optimization, Genetic Algorithm and Pattern Search (PS) respectively. Both Nelder-Mead Simplex and Pattern Search are categorized in local optimization method. However, the algorithm differs to each other. In Pattern Search method, the objective function is updated around their neighborhood points, while Nelder-Mead Simplex method updating their objective function by comparing the value at the three vertices of a triangle. Besides that, the experimental data used in this study is different from [22]. Hence we have surviving fraction cell of HeLa cell which are treated with radio sensitizer. There are evidences showing that the used of radio sensitizer increase the number of cell killing in radio therapy [23,24].

The aim of this method is to minimize the sum of square error (SSE) between experimental data and simulation data. The performances of both algorithms are compared based on the value of SSE and computational time.

3 Optimization Procedure

3.1 Genetic Algorithm

Genetic Algorithm was first proposed by John Holland [25]. This method generates solutions to optimize problem using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. Besides that, Genetic Algorithm is an adaptive heuristic method and successfully applied to solve the optimization problem. It is a reliable and accurate method of an optimization problem with a series of natural process involved [26]. For example, Lee et al. [27] applied Genetic Algorithm to estimate parameter in the mesoscale meteorological model in order to improve the quantitative precipitation for heavy rainfall case in Korea [27] as well as Martín-Cañada et al. who used Genetic Algorithm to estimate parameter in the neuron model [28].

The solution generated by this method is called chromosome. The chromosome is represented as a set of parameters that defines the solution to the problem. Gene is joined to form a chromosome, while population refers to a group of the chromosome. Its value is in binary string, 0 and 1.

Crossover is the most important part in Genetic Algorithm where the offspring is created. It involves two individual parents mating with each other to produce two new offsprings. The process keeps repeating in order to generate a high quality offspring. The process terminates after the value converged to a certain value. The basic terminologies in Genetic Algorithm, as well as the main step to estimate the parameter for every model can be presented in Figure 1 and summarized as follows:

1. Generate initial population for individual’s chromosome (parameters).
2. Evaluate the fitness of each chromosome.
3. Create new chromosome (offspring) by mating current chromosome using a suitable operator. There are four operators of Genetic Algorithm:
   i. Selection
   ii. Crossover
   iii. Mutation
iv. Reproduction

4. Place the new chromosome in the new population

![Figure 1: The Framework in Optimizing using Genetic Algorithm](image)

### 3.2 Pattern Search

Pattern Search (PS) also known as Hooke-Jeeves method is a free derivatives method in searching minimum value [28]. This method works by creating a set of search direction iteratively. It relies on the value of \( f(x) \) on a sequence \( x_1, \ldots, x_n \). The value of objective function, \( f(x) \), is compared to search the minimum value. In 2003, Honkela employed Pattern Search for parameter estimation [30]. Later, Emary et al. [31] use Pattern Search to optimize the segmentation of retinal vessels based on multi-objective segmentation [31].

It involves two types of move: exploratory move and pattern move. Exploratory move is the initial move that looks for an improving direction to move. This is done by perturbing the current point by a small amount in each variable direction and the objective value is observed. Pattern move require two points, the current point and another point that has a better value of objective function. This will give the pattern move an improving direction to move in.

The method proceeds by a sequence of exploratory and pattern move from a base point. If an exploratory move leads to a decrease in the value of \( f(x) \) it is called a success, otherwise, it is called a failure. The algorithm for pattern search method as well as the flowchart is shown in Figure 2 and summarized below:

**Step 1:** Selection of initial condition which includes:

i. Initial points: \( x(0) \)
ii. Increment vector: \( \Delta i = (0.5, ..., N) \)

iii. Reduction vector: \( \alpha = 2 \)

iv. Termination parameter: \( \varepsilon = 10^{-3} \)

Step 2: Perform the iteration with the starting point, \( x(0) \). Three points are set as the function value. Calculate

i. \( f_+ = f(x + \Delta) \)

ii. \( f = f(x) \)

iii. \( f_- = f(x - \Delta) \)

Step 3: Find \( f_{\min} = \min(f_+, f, f_-) \) if \( f_{\min} < f(x) \) then \( f(x) \) is replaced by \( f_{\min} \) otherwise, retain \( f(x) \).

Step 4: Repeat step 2 for the second iteration until the minimum value of the function is reached.

---

3.3 Parameter Estimation Procedures using GA and PS

3.3.1 Genetic Algorithm

Step 1: Set the initial parameters value.

Step 2: Based on the initial value in Step 1, the objective function (SSE) is calculated by using GA technique.

Step 3: Estimated parameters value are obtained.
3.3.2 Pattern Search

Step 1: Set the initial parameters value.

Step 2: Based on the initial value in Step 1, the objective function (SSE) is calculated by using PS technique.

Step 3: Estimated parameters value are obtained.

Both results are then compared. The comparison between these two optimization techniques are based on the value of the objective function, SSE and computational time.

4 Results and Discussion

An experimental data from [32] is utilized for the purpose of parameter estimation. The experimental data is a treatment of HeLa cell treated with 29 nm of platinum nanodentride (PtNDs), a type of radiosensitizer and the cells is irradiated with a 6 MV photon beam for this purpose. It is a compound used to increase the effect of ionizing radiation. There are evidences show that the use of radiosensitizer can cause double-strand break (DSB), the most lethal DNA damage [33,34]. The experimental data points are plotted in Figure 3. The experimental data is fitted to the LQ formulation \( \ln S = -0.0593D - 0.0342D^2 \) where \( \alpha_{\text{exp}} = 0.0593 \) and \( \beta = 0.0342 \) using GA method.

![Figure 3: The Survival Curve of HeLa Cell treated with 29nm of Platinum Nanodentride (PtNDs), Combined with a 6 MV Photon Beam](image-url)
The experimental data in Figure 3 is fitted to the LQ formulation \( \ln S = -0.0593D - 0.0342D^2 \) where \( \alpha_{\text{exp}} = 0.0593 \) and \( \beta = 0.0342 \) using GA method. In this study, the initial value of six parameters of the model \( \delta, \alpha_1, \alpha_2, \rho, V_{\text{max}}, K_M \) are set randomly. The description of each parameter is shown in Table 1.

In this study, the built-in optimization toolbox in MATLAB is used, which are ga for Genetic Algorithm and patternsearch for Pattern Search. The performance of both optimization methods are evaluated based on the value of SSE and computational time. The results of parameter estimation for 50 runs are shown in Table 2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>( \alpha_{\text{exp}} )</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \delta )</td>
<td>Radiosensitivity of cell</td>
<td>0.0593</td>
<td>0.0342</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>Death rate due to misrepair DSB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>Death rate due to lethal aberration of DSBs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho )</td>
<td>Probability of correctly repair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( V_{\text{max}} )</td>
<td>Maximum repair rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( K_M )</td>
<td>Michaelis-Menten constant</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 presents the mean of estimated parameter value by GA and PS. From the table, it can be seen that SSE value for GA is smaller than PS value. The value of SSE for GA is close
to 0 which correspond to an excellent fit between model and experimental data. The value of the standard deviation of each parameter is also provided in the table. Standard deviation is a measure of variation about the mean of parameters value. It is observed, the large value of standard deviation indicates more spread out of data [35]. In this study, the value of the sample standard deviation of the estimated value for $\delta, \alpha_1, V_{\text{max}}$, and $K_M$ given by GA are larger than the other parameter. This means that these parameters are insensitive to the model.

From the result, it is found that GA is more superior than PS method in terms of in providing the smallest value of SSE. This result is supported by the computational of correlation, $r$, between real data and simulation data provided by GA which is close to 1. The value which is close to 1 indicates an excellent fit between model and the experimental data. Figure 4 provides the plot between experimental and simulation data, which explain the good fit between experimental and simulation data.

The computational time is also taken into consideration in order to compare the performance of both optimizers. The time refers to the time taken for optimizer methods to converge to the objective function to a minimum value. The time taken for both optimizers are presented in Table 3.

![Figure 4: The Computational Time in Minimizing the Objective Function (SSE) using GA and PS](image)
Table 3: The Computational Time in Minimizing the Objective Function (SSE)

<table>
<thead>
<tr>
<th>Method</th>
<th>Time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
<td>684.2272</td>
</tr>
<tr>
<td>PS</td>
<td>528.852</td>
</tr>
</tbody>
</table>

From Table 3, it shows that the time taken for GA to minimize the objective function is longer than PS method. As discussed in Section 4, GA algorithm is more complicated compared to PS method. Genetic Algorithm generates a solution based on the chromosome culture. It involves two parents chromosomes in creating two new chromosomes. The process ends after the value of the selected gene in the chromosome changed from 1 to 0 and vice versa. Therefore, the process consumes a long time to converge to a minimum value. However, GA manages to obtain the smallest $SSE$ value which is 0.00055. Thus, it is concluded that in our case, GA is more superior than PS.

Next, the computational of 95% confidence interval (CI) of each parameter in the model are provided. The confidence interval gives the interval value of each parameter around its mean value. In this study, we use 95% CI which means that it is 95% true the value of parameters is within the interval. The confidence interval for GA is computed and shown in Table 4.

Table 4: The 95% Confidence Interval for the Six Parameters of the Model using GA Method

<table>
<thead>
<tr>
<th>Method</th>
<th>Parameter</th>
<th>Confidence Interval (C.I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
<td>$\delta$</td>
<td>(2.7654,3.3477)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>(6.1534,9.9057)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_2$</td>
<td>(0.0015,0.0023)</td>
</tr>
<tr>
<td></td>
<td>$\rho$</td>
<td>(0.9769,0.9829)</td>
</tr>
<tr>
<td></td>
<td>$V_{max}$</td>
<td>(0.9644,1.4242)</td>
</tr>
<tr>
<td></td>
<td>$K_M$</td>
<td>(3.3582,4.1497)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{\text{model}}$</td>
<td>(0.0472,0.0545)</td>
</tr>
<tr>
<td></td>
<td>$\beta_{\text{model}}$</td>
<td>(0.0433,0.0475)</td>
</tr>
</tbody>
</table>

5 Conclusion

The aim of this study is to estimate parameter’s value in the model using survival data of HeLa cell treated with a combination of radiation and radiosensitizer. The used of radiosensitizer in radiotherapy is to enhance the number of cell killing. Parameter estimation is carried out to determine the descriptive measure of cell response toward radiosensitizer and radiation dose. Local and global optimization have been used for this purpose. The results of parameter estimation between GA and PS are compared to determine which method is more efficient for this study. The results showed that, GA provides the smallest value SSE compared to PS. On the other hand, the computational time consumed by GA is longer than PS method due to
the complexity of GA algorithm. However, GA managed to get the smallest SSE value despite the longer computational time. Besides that, the correlation between experimental data and simulation provided by GA is close to 1 which indicates an excellent fit.

Acknowledgement
This work was financially supported by the Universiti Teknologi Malaysia and under the Fundamental Research Grant Scheme (FRGS: R.J130000.7826.4F889) and GE STEM grant with no. 07397 [Ministry of Education Malaysia (MOE)]. The first author would like to express her appreciation to MOE for her MyPhD Scholarship.

References


[29] Hooke, R. and Jeeves, T. A. Direct search solution of numerical and statistical problems. 


