Statistical and Nature-inspired Metaheuristics Analysis on Flexirubin Production

Siti Nurulasilah Suhaimi[1,2], Shafaatunnur Hasan[1,2*], Siti Mariyam Shamsuddin[1,2], Wan Azlina Ahmad[3], Chidambaram Kulandaisamy Venil[3]

[1]UTM Big Data Centre, Universiti Teknologi Malaysia,
81310 Johor Bahru, Johor, Malaysia
[2]Faculty of Computing, Universiti Teknologi Malaysia,
81310 Johor Bahru, Johor, Malaysia
e-mail: siti.nurulasilah.suhaimi@gmail.com, shafaatunnur@utm.my,
mariyam@utm.my
[3]Faculty of Science, Universiti Teknologi Malaysia,
81310 Johor Bahru, Johor, Malaysia
e-mail: azlina@kimia.fs.utm.my, ckvenil@gmail.com

Abstract

Nowadays, demand for natural pigments has increased dramatically due to the awareness of the toxicity of some synthetic pigments. Because of the high cost of growth medium for natural pigment production, various studies have been carried out to explore medium which are less costly, such as agricultural waste. This study highlight on the application of firefly algorithm (FA) and bat algorithm (BA) in optimizing yellowish-orange pigment production (flexirubin) from the agricultural waste material. At present, response surface methodology (RSM) is the most preferred statistical method in optimizing pigment production. However, in the last two decades, nature-inspired metaheuristics approach has been used extensively in the fermentation process and have continually improve the efficiency in the optimization problem especially in pigment production. This study compared the analytics studies of RSM, FA and BA in the estimation of fermentation parameters (Lactose, L-tryptophan, and KH_2PO_4) in flexirubin production from Chryseobacterium artocarpi CECT8497^T. All models provided similar quality predictions for the above three independent variables in term of flexirubin production with bat algorithm showing more accurate in estimation, with the coefficient value of 98.87% compare to RSM 98.20% and FA 98.38%.

Keywords: *Agricultural waste, Bat algorithm, Firefly algorithm, Optimization, Response Surface Methodology*

1   Introduction

Pigment is a substance used as coloring [1]. Pigment involves in most of the industries nowadays, including food production, textile industries, cosmetics industries, painting and many more. They used pigment or colorant in order to make the products appear colorful and attractive to consumers, thereby increasing their profit margins. Nowadays, consumers have been aware to toxicity of some synthetic pigments that can cause a hazard to human health and the environment. Therefore, pigment from natural sources has gained tremendous demand and has been increased dramatically.

The major obstacles of producing organic pigments normally hampered by the high production cost. Therefore, it is important to find an alternative pigment which can be biodegraded and easily available at a minimal production cost, such as agricultural waste. Besides that, the use of these agro-industrial residues in the pigments bio-process can also reduce its environmental impact [1]. There are various organisms that live in the agricultural wastes such as microorganisms. Some microorganisms that thrive in these agricultural wastes have the ability to produce pigment. In this study, *Chryseobacterium artocarpi* collected from the agricultural waste residue has been used to produce a pigment called flexirubin (yellowish-orange pigment).

In the fermentation process, medium optimization is recognized as a simple but effective method for achieving high productivity of desired product [2]. The most important thing in efficient utilization fermentation technology and pigment production process is the development of proper fermentation medium [3]. The use of different growth medium, directly affect the growth and production of pigments. Nowadays, response surface methodology (RSM) known as statistical technique is the most informative method and have been used extensively in medium fermentation optimization. However, due to the high demand for a high pigment production, various optimization algorithms have been studied to enhance the modelling and optimization of the fermentation medium. For several decades, nature-inspired computing have gained popularity in optimization research, due to the capability to solve a variety of problems and the successful application in complex optimization problems, such as firefly algorithm (FA) and bat algorithm (BA).

FA and BA belongs to swarm intelligence group, which is characterized as a biological system that have been influenced by collective behaviors of nature like swarms or insect colonies. Firefly algorithm [4] was developed by Xin She Yang in 2008, based on the flashing patterns and behavior of fireflies. Various versions of FA algorithms were used in many different problems such as optimization of glucose concentration [5], power scheduling [6], electrical discharge [7], and robot workcell layout [8]. Bat algorithm [9] was developed by Xin She Yang in 2010, based on the echolocation behavior of bats. There are various optimization

problem that used BA such as optimization of power system stabilizer [10], job shop scheduling [11], and optimization for an autonomous mobile robot [12].

Due to the robustness and flexibility of FA and BA techniques in optimizing problems, this study was intended to evaluate the computational effectiveness of BA and FA for the optimization of flexirubin production. This paper reports on the analysis studied between three optimization techniques – RSM, FA and BA were used to enhance the flexirubin production by *Chryseobacterium artocarpi* CECT 8497. These three methodologies were compared for their predictive ability, sensitivity analysis and optimization capabilities. The predictive ability of statistical technique and swarm intelligence were compared using dataset of 17 experiments.

2   Related Work

There are limited number of studied in optimizing natural pigment production using metaheuristics algorithm. Currently, artificial neural network (ANN) and genetic algorithm (GA) are usually used for the optimization of fermentation medium in natural pigment production. ANN have gained wide popularity in the decision making and are being applied in many optimization of pigment production. Singh et al. [16] studied the application of ANN in optimizing red pigment production using *Monuscus purpureus* bacteria, while Sinha et al. [17] studied the optimization of microwave-assisted extraction of yellow-red natural color from Annatto seed. Both result shows that ANN can predict the effect of fermentation parameters on pigment production with a high correlation. Apart from that, Joshi and Singhal [18] used hybrid ANN-GA in optimizing zeaxanthin production by *Paracoccus zeaxanthinifaciens* ATSS 21588 in a complex medium. Zeaxanthin is an orange-red colored. From the results, ANN-GA hybrid performed better prediction efficiency and enhanced the zeaxanthin production by 21%. Hence, it shows that metaheuristics algorithm have better prediction efficiency in optimizing pigment production. At present there are no reports in optimizing flexirubin production using FA and BA based techniques.

3   Research Framework
3.1   Data set preparation and parameter setting

In this study, a bacterium known *Chryseobacterium artocarpi* CECT 8497T (= KCTC 32509T) isolated from agricultural waste in the orchard in Universiti Teknologi Malaysia (UTM), Skudai, Malaysia has been used [14]. This strain produce pigment-type known as flexirubin and it is identified as yellowish-orange color [14, 15]. This strain was cultivated in 100 mL Erlenmeyer flask containing 20 mL of nutrient broth (NB). After that, the flask was incubated at 30^0C with a shaking speed of 200 rpm for 24 hour. Next, the remaining sample was extracted

using acetone. Next the sample have been sonicate, filter and quantified for pigment yield (flexirubin).

In this experiment, three types of nutrient broth (independent variable) were used; L-tryptophan, lactose and potassium dihydrogen phosphate (KH_2PO_4), respectively. The levels of the chosen independent variables used in the experiments are given in Table 1, where the effect of different parameters such as lactose concentration (10-12.5 mg/L), L-tryptophan (4-8 mg/L) and KH_2PO_4 (0.3-1 g/L) on the flexirubin production (mg/L).

Table 1: Independent parameters and experimental design levels

Independent parameters	Symbol	Actual levels of code factors		
		-1	0	+1
Lactose (mg/L)	X_1	10	11.25	12.5
L-tryptophan (mg/L)	X_2	4	6.0	8
KH_2PO_4 (g/L)	X_3	0.3	650	1

Next, box-behnken design (BBD) was used to optimize the response of the independent variables on flexirubin production (response). Three independent parameters, Lactose (x_1), L-tryptophan (x_2) and KH_2PO_4 (x_3) have been randomized according to box-behnken design, where the design contains 17 experiment trials as in Table 2. As presented in Table 2 (No. 6), a moderately lactose concentration of 11.25 mg/L, L-tryptophan of 6.00 mg/L and KH_2PO_4 of 0.65 g/L cause maximum flexirubin production, and it is verified when the optimum values of nutrient concentration for the flexirubin production is 501.48 mg/L, while the predicted value of RSM is 498.15 mg/L. The experimental and RSM predicted data in Table 2 has been taken from previous paper from Venil et al. [14].

Table 2: Experimental and predicted values (RSM)

No	Factors			Experimental (mg/L)	RSM predicted (mg/L)
	X_1	X_2	X_3		
1	+1	0	+1	172.22	169.24
2	-1	-1	0	212.98	224.10
3	0	-1	+1	350.18	327.19
4	0	0	0	480.51	478.14
5	0	-1	-1	378.40	333.29
6	0	0	0	501.48	498.15
7	+1	-1	0	203.46	198.47
8	0	0	0	483.98	490.43
9	-1	0	-1	199.12	209.05

10	+1	+1	0	170.39	181.47
11	+1	0	-1	249.21	255.42
12	-1	0	+1	151.53	162.63
13	0	0	0	472.51	467.42
14	0	+1	-1	297.54	280.51
15	0	0	0	483.11	491.30
16	-1	+1	0	194.20	197.02
17	0	+1	+1	252.95	268.41

3.2   Modelling and optimization for pigment production

The optimization of fermentation conditions and medium composition is of primary importance in any process owing to their impact on the economy and practicability of the process [13]. Therefore in this study, metaheuristics algorithm (FA and BA) have been applied in the fermentation process to gain the optimal design of medium composition, hence to improve the production of *flexirubin*.

Fig. 1 shows schematic representation of metaheuristics implementation in fermentation process, which the left-hand side represent the wet-lab experiment condition process while the right-hand side demonstrated the metaheuristics algorithm process. The explanation for left-hand side of flowchart has been described in section 3.1. As can be seen from right-hand side of Fig. 1, metaheuristics optimization have been applied in fermentation culture process. Before the implementation of metaheuristics algorithm, multiple regression model was developed to obtain the objective function for the metaheuristics algorithm. General step for the implementation of metaheuristics algorithm has been shown in Fig. 1. In this process, FA and BA algorithms have been used to optimize the medium composition of nutrient broth for *Chryseobacterium Artocarpi* bacterium to growth. Further explanation for the implementation of FA and BA algorithm in the fermentation culture process are being discussed in next sub-section.
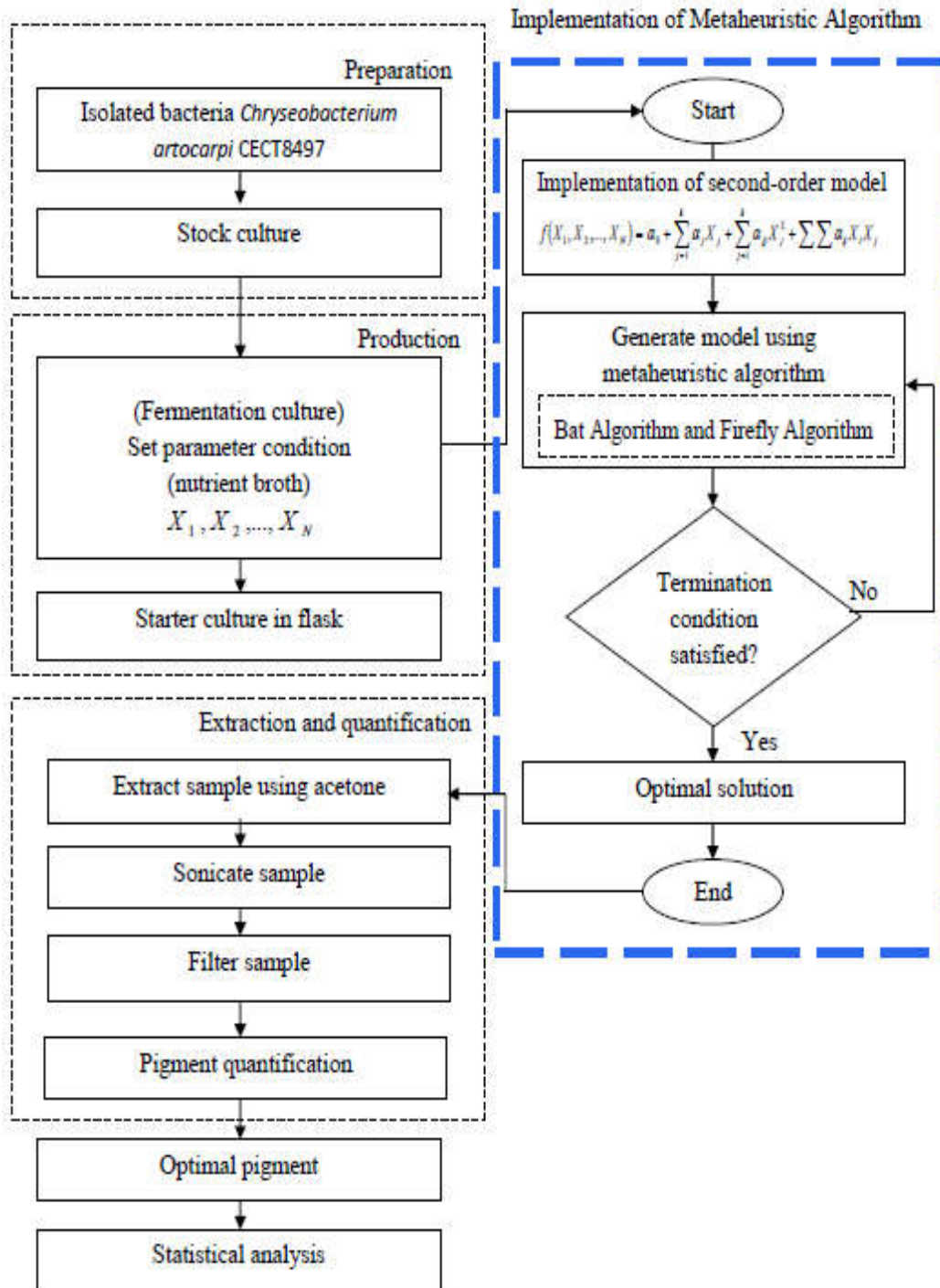
Implementation of Metaheuristic Algorithm

Preparation

Isolated bacteria *Chryseobacterium artocarpi* CECT8497

Stock culture

Production

(Fermentation culture)
Set parameter condition
(nutrient broth)
$X_1, X_2, ..., X_N$

Starter culture in flask

Extraction and quantification

Extract sample using acetone

Sonicate sample

Filter sample

Pigment quantification

Optimal pigment

Statistical analysis

Start

Implementation of second-order model

$f(X_1, X_2,..., X_N) = a_0 + \sum_{j=1}^{k} a_j X_j + \sum_{j=1}^{k} a_{jj} X_j^2 + \sum \sum a_{ij} X_i X_j$

Generate model using metaheuristic algorithm

Bat Algorithm and Firefly Algorithm

Termination condition satisfied?

No

Yes

Optimal solution

End

Fig.1: Schematic representation of Metaheuristics Implementation in
Fermentation Process

3.2.1 Multiple regression model for FA and BA objective function

Multiple regression model are developed based on the experimental data given in Table 2. The data were subjected to the analysis of variance (ANOVA) and fitted accordingly to second-order model. A mathematical model, describing the relationship between the dependent variable and independent variables (lactose, L-tryptophan and KH_2PO_4) in a second-order polynomial model was developed accordingly as given in Eq. (1).

$$f(x) = \alpha_0 + \sum_{j=1}^{k} \alpha_j X_j + \sum_{j=1}^{k} \alpha_{jj} X_j^2 + \sum \sum \alpha_{ij} X_i X_j \tag{1}$$

where $f(x)$ is the predicted response, X_i and X_j are coded as independent variables, while $\alpha_0, \alpha_j, \alpha_{jj}, \alpha_{ij}$ are the regression coefficient for the intercept, linear, quadratic and interaction effect, respectively.

$$f(x) = -17526.123 + 3017.057\ x_1 + (248.975)\ x_2 + (1038.580)x_3 +$$
$$(-133.059)x_1^2 + (-20.289)x_2^2 + (-680.767)x_3^2 + \tag{2}$$
$$(-1.429)x_1x_2 + (-16.800)x_1x_3 + (-5.846)x_2x_3$$

where $f(x)$ is yield (*flexirubin* production), x_1, x_2 and x_3 represent lactose, L-tryptophan and KH_2PO_4, respectively.

Table 3: Multiple linear regression analysis of the interaction between variables

Parameters	Coefficients	t-value	P-value
Constant	-17526.128	-16.968	<0.05
X_1	3017.057	17.334	0.000
X_2	248.975	3.755	0.007
X_3	1038.580	2.940	0.022
$X_1^{\,2}$	-133.059	-17.559	0.000
$X_2^{\,2}$	-20.289	-6.854	0.000
$X_3^{\,2}$	-680.707	-7.043	0.000
$X_1 X_2$	-1.429	-0.294	0.777
$X_1 X_3$	-16.800	-0.605	0.564
$X_2 X_3$	-5.846	-0.337	0.746

Table 3 shows the results of the analysis of variance (ANOVA) of the regression equation model. As seen in Table 3, the results of the model showed that the p-value of the model terms were significant, p <0.05, except interactions of lactose with L-tryptophan, lactose with KH_2PO_4 and L-tryptophan with KH_2PO_4. The p-value indicates whether a term in a model is significant or not, and the Fisher test (F-value) shows the level of significance in the model terms. The model Fisher test F-value of 51.141 with corresponding low probability p-value (p<0.05) imply the model obtained is statistically highly significant for pigment production. The fit model was also expressed with the correlation coefficient, r^2 it is used as a tool to check the model and it said perfect fit when the value near to unity. As presented in Fig. 2, r^2 of the model obtained was 0.9850, which indicates that the sample variation of 98.50% for flexirubin concentration is attributed to the independent factors and only 1.50% of the total variation are not explained by the model. The paired-sample t test has demonstrated the positive correlation by reducing mean value from the wet-lab experimental results by 309.0453, $t(15) = 0.340$. From the laboratory analysis, together with statistical tests revealed that the equation, second-order polynomial equation (Eq. 2) was being utilized as the objective function for FA and BA.

3.2.2   Firefly Algorithm

Firefly algorithm (FA) [4] is a nature inspired metaheuristic algorithm, developed by Xin-She Yang in 2008, inspired by the flashing behavior of fireflies, which search for a food and mating partner by emitting a flashing light. There are about two thousand firefly species, and most fireflies produce short and rhythmic flashes or well known as bioluminescence. The primary purpose for a firefly's flash are to attract mating partners (communication), attract potential prey and serve as a protective warning mechanism.

> The basic assumptions in Firefly Algorithm are
1) All fireflies are unisex, where every firefly can be attracted to other fireflies irrespective of their sex.
2) Attractiveness is proportional to their light intensity, thus the less attractiveness fireflies will move towards the brighter one. Attractiveness and brightness both decrease as their distance increases. If there is no one brighter than a particular firefly, it will move randomly by random walks
3) The brightness of a firefly is determined by the view of the objective function.

The initial population of fireflies is generated randomly and by using the fitness function of the optimization problem, the attractiveness β_0 is determined for each firefly in the population. All fireflies move throughout the solution space for a certain number of iterations. For every iteration, the attractiveness of each two

fireflies f_i and f_j is compared, if f_i is more attractive than f_j then firefly f_j will move toward firefly f_i. In the firefly algorithm, there are two important points, which are the variation in the light intensity and formulation of the attractiveness. However, the light intensity I and attractiveness β depend on the distance r between fireflies and the media light absorption coefficient γ.

In this study, it is assumed that the attractiveness of the firefly is determined by its brightness, which is associated with the objective function $(f(x))$ of the optimization of *flexirubin* production. Each firefly has a location $X = (X_1, ... X_d)$ in d-dimensional space, and the light intensity and attractiveness are proportional to an objective function $(f(x))$. In this case, $f(x)$ is equal to Eq. (2).

The light intensity, I varies with distance r is expressed by the following equation

$$I(r) = I_0 e^{-yr^2} \qquad (3)$$

where I_0 denotes the intensity of the light source, and γ is the fixed light absorption coefficient. Similarly, in the equation below the attractiveness β also depends on the distance r

$$\beta(r) = \beta_0 e^{-yr^2} \qquad (4)$$

where r is the distance between any two fireflies, β_0 is the attractiveness at $r = 0$.

The distance between any two fireflies i and j at x_i and x_j, respectively, is the Cartesian distance as follows:

$$r_{ij} = |x_i - x_j| = \sqrt{\sum_{k=1}^{D}(x_{i,k} - x_{j,k})^2} \qquad (5)$$

where $x_{i,k}$, is the k-th element of the i-th firefly position within the search-space, and denotes the dimensionality of a problem.

The movement of a firefly i is attracted to another more attractive (brighter) firefly j, is established by the equation below

$$X_i = x_i + \beta_0 e^{-yr^2}(x_j - x_i) + \alpha \times (rand) \qquad (6)$$

The following equation consists of three terms, where the first term determines the position of the i-th firefly. The second term is due to the attraction while the

third term is randomization with α being the randomization parameter, and the random number $\in$ drawn from a Gaussian distribution. Fig. 2 demonstrates the representation of firefly movement in which each firefly represents a candidate solution ($X_1, X_2, and\ X_3$), in case of study of *flexirubin* production.
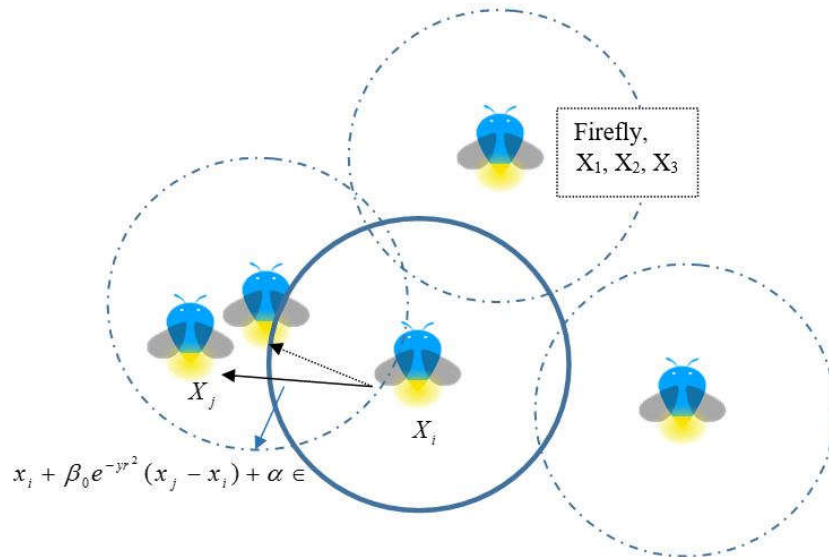

Fig.2: Graphical representation of firefly movement

Fig. 3 demonstrated the schematic diagram of representation FA implementation in fermentation process, which the left-hand side represent the wet-lab experiment condition process while the right-hand side demonstrated the implementation of BA in fermentation culture process. The explanation for left-hand side flowchart has been described in section 3.1. While, the right-hand side of the flowchart demonstrated the process of FA. For a maximization problem, the brightness is simply proportional to the value of the objective function. In FA algorithm, parameter values used were $\alpha \in [0, 1]$, $\beta_0 = 1$, $n = 50$, and the number of generations was set to 500.
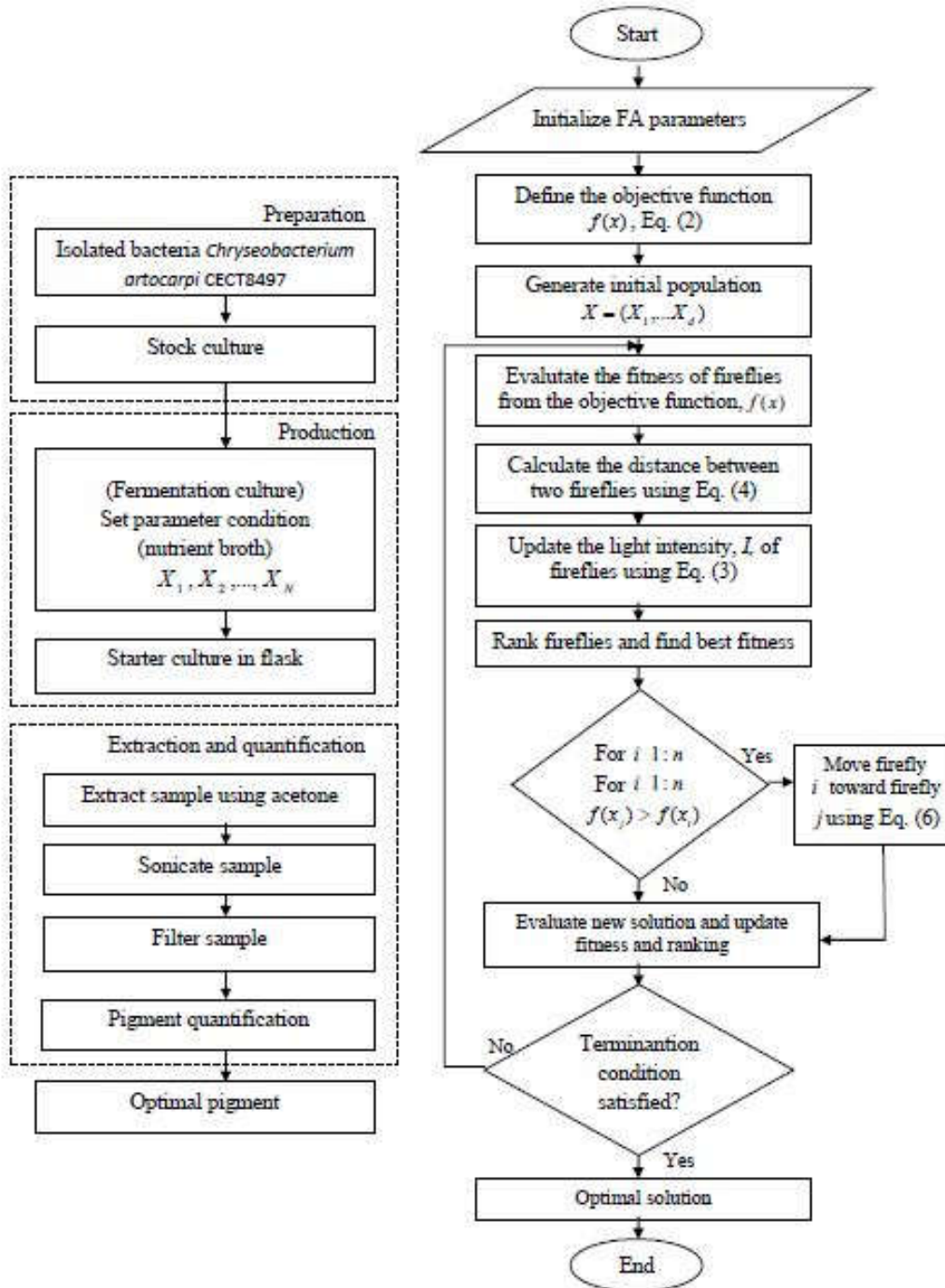
Fig.3: Schematic representation of Firefly Algorithm Implementation in Fermentation Process

3.2.3 Bat Algorithm

Bat algorithm [9] was proposed by Xin She Yang in 2010, and inspired by the echolocation behavior of bats. Bats are fascinating animals, where bats was the only mammals with wings and have innovative capability of echolocation. Echolocation is a type of sonar used by microbats to locate object, avoid obstacle, detect prey, and pinpoint prey by reflected sound [19]. Some bats have evolved a highly sophisticated sense of hearing and it depends on the species. Bats emit a high sound frequency to neighborhood and hear the echo that bounces back to the bats. From these echoes, bats can determine the size of the objects, how far the objects, and texture of the objects.

Yang idealized some of the echolocation characteristics of microbats and developed exotic algorithm known as bat algorithm. Echolocation characteristics of microbats emphasize on some of approximate rules

(a) All bats use echolocation to sense distance. They acknowledge the difference between food/prey and surrounded barriers in some miraculous ways.

(b) Bats fly randomly with velocity at position with a fixed frequency, varying   wavelength λ and loudness to search for prey. They can automatically adjust the   wavelength of their emitted pulses and adjust the rate of pulse emission in the range of [0, 1], depending on the proximity of their target.

(c) Although loudness can vary in many ways, The loudness varies from large (positive) A_0 to minimum constant value $A_{\min}$.

Initial population is generated randomly by considering lower and upper boundaries with the dimension d and the number of bats N.

$$x_{ij} = x_{\min j} + (x_{\max j} - x_{\min j}) \times rand \tag{7}$$

where $i = 1,2,...,n$, $j = 1,2,...n$, $rand$ is a random vector [0,1], $x_{\max j}$ and $x_{\min j}$ represent upper and lower boundaries, respectively.

Initially, frequency of bat algorithm is set random value for each solution in range of $f_{\min}$ and $f_{\max}$. Then new solution x_i^t and velocity v_i^t are updated using the following equation;

$$f_i = f_{\min} + (f_{\max} - f_{\min}) \times rand \tag{8}$$

$$v_i^t = v_i^{t-1} + (x_i^t - x_{Gbest}) f_i \tag{9}$$

$$x_i^t = x_i^{t-1} + v_i^t \tag{10}$$

where x_{Gbest} represents global best solution. Once a solution is selected among the best solution for the local search, a new solution is generated locally using random walk

$$x_{new} = x_{Gbest} + \varepsilon A^t \tag{11}$$

where $\varepsilon \in [-1,1]$ is a uniform random number, while A^t is the average loudness of all the bats.

Loudness A_i and the rate r_i of pulse emission have to update as iterations proceed. When the bat gets closer to its prey, the loudness decreases while pulse emission increases.

$$A_i^{t+1} = \alpha A_i^t \tag{12}$$
$$r_i^{t+1} = r_i^0 [1 - \exp(-\gamma t)] \tag{13}$$

where α and γ are constants. Normally, each bats have different values of loudness [1, 2] and pulse emission rate [0, 1], and this can be achieved by randomization. After these bats are moving towards the optimal solution or new solutions are improved, their loudness and pulse emission rate will be updated. Fig. 4 demonstrates the representation of bat movement in which each bat represents a candidate solution ($X_1, X_2, and\ X_3$) to a given problem (in case to optimize the response of variables on *flexirubin* production).
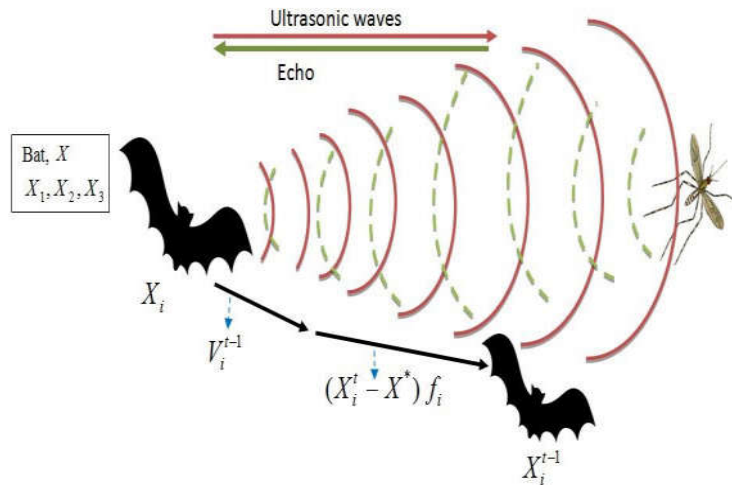
Fig.4: Candidate solution movement in Bat Algorithm

Fig. 5 demonstrated the schematic diagram of representation of BA implementation in fermentation process, which the left-hand side represent the wet-lab experiment condition process while the right-hand side demonstrated the implementation of BA in fermentation culture process. The explanation for left-

hand side flowchart has been described in section 3.1. While, the right-hand side of the flowchart demonstrated the process of BA. For BA algorithm, several parameters must be set to initial values. In this study the population size is set to $50 (N = 50)$, dimension was set to 30, the random numbers are between [0, 1], and the number of generations was set to 500.
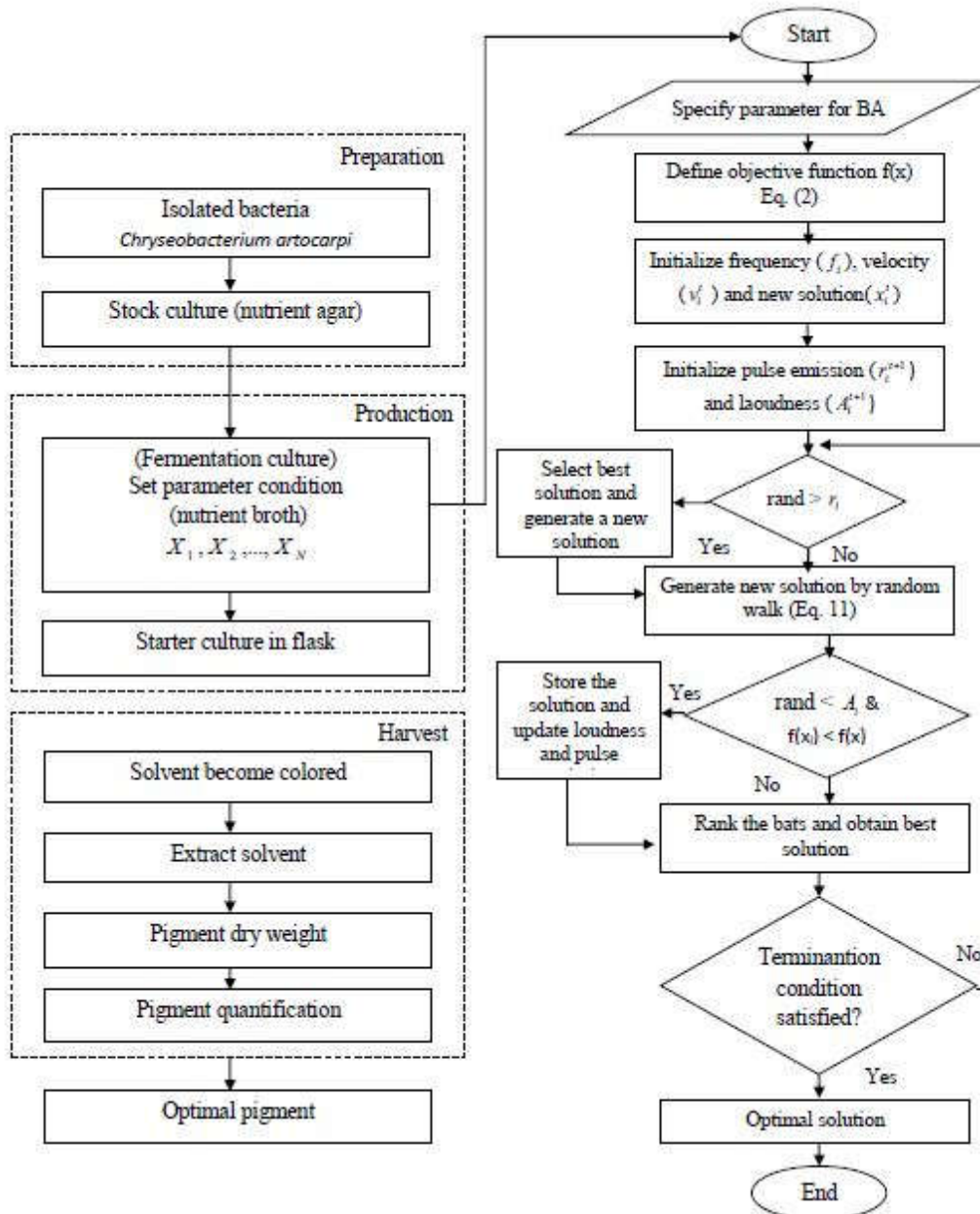
Fig. 5: Schematic representation of Bat Algorithm Implementation in Fermentation Process

4   Result and Analysis

Previously, the production of flexirubin by bacteria strain *C. artocarpi* CECT 8497T (=KCTC 32509T) has been optimized using RSM by Venil et al. [14]. The enzyme production was significantly improved by three medium components, including lactose (X_1), L-tryptophan (X_2) and KH₂PO₄ (X_3), as presented in Table 4. In this section, RSM was compared with other well-known algorithm as FA and BA. The predicted and actual experiments of flexirubin production are presented in Table 4.

As seen in Table 4, the measured and predicted flexirubin production with swarm intelligence (BA and FA) techniques were almost identical as compared to the predicted value by the RSM. Though RSM models performed well in terms of the r^2, and offered stable responses, yet the FA and BA techniques was better in both data fitting and estimation capabilities in comparison to RSM. This was also corroborated by the scatter plots between the predicted values (BA and FA) and experimental values as presented in Fig. 6.

Table 4: Experimental design for RSM, FA and BA of flexirubin production by *C. artocarpi* CECT 8497T

| Run | Factors | | | Yellowish-Orange Pigment (mg/L) | | | |
| | X_1, | X_2, | X_3, | | Predicted | | |
	(mg/L)	(mg/L)	(ppm)	Experimental	RSM	FA	BA
1	+1	0	+1	172.22	169.24	178.44	170.67
2	-1	-1	0	212.98	224.10	217.96	218.50
3	0	-1	+1	350.18	327.19	340.01	340.33
4	0	0	0	480.51	478.14	487.76	487.32
5	0	-1	-1	378.40	333.29	369.43	372.73
6	0	0	0	501.48	498.15	487.32	487.76
7	+1	-1	0	203.46	198.47	236.29	230.30
8	0	0	0	483.98	490.43	487.76	487.32
9	-1	0	-1	199.12	209.05	204.69	200.93
10	+1	+1	0	170.39	181.47	172.40	169.64
11	+1	0	-1	249.21	255.42	230.39	233.47
12	-1	0	+1	151.53	162.63	181.80	175.56
13	0	0	0	472.51	467.42	487.76	487.32
14	0	+1	-1	297.54	280.51	320.88	318.37
15	0	0	0	483.11	491.30	487.76	487.32
16	-1	+1	0	194.20	197.02	168.39	172.57
17	0	+1	+1	252.95	268.41	275.19	267.16
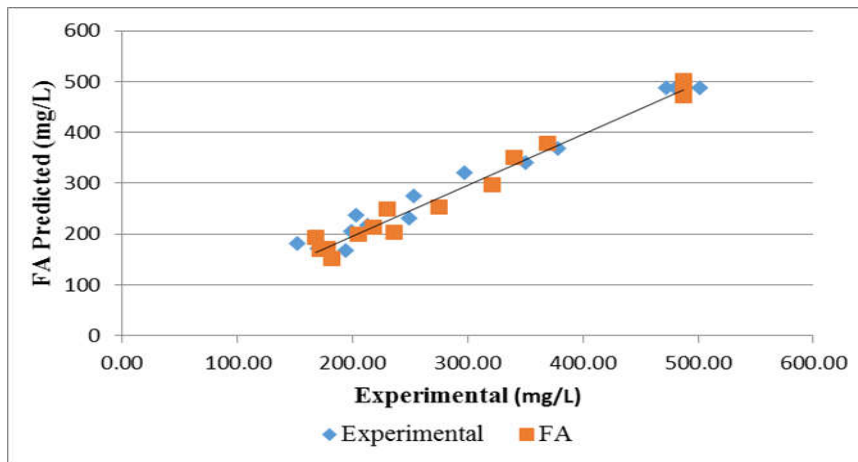
(A)
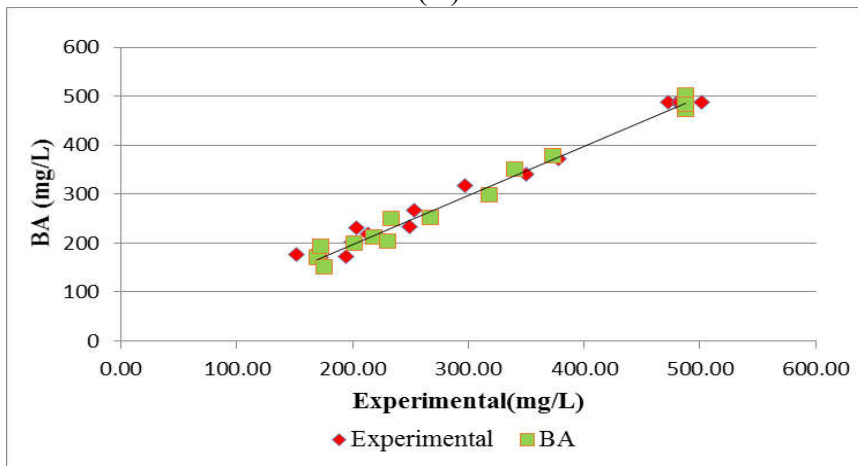
(B)

Fig. 6: (A-B) Scatter plot of model predicted using FA and BA versus experimental.

A scatter plot of predicted firefly algorithm (FA) against actual experiment for flexirubin production was shown in Fig. 6(A). In this study, r^2 value for FA model was 0.9838, demonstrated that there was 98.38% of the variability in the response can be explained by the model. Fig. 6 (B) shows a scatter plot of the data predicted bat algorithm (BA) against actual experiment for optimization of the flexirubin production. The plots showed that there is significant interaction among the variables, it can be verified from the linear regression when r^2 was found approximately 0.9887 indicated that 98.87% of the variability in the response can be explained by the model It exemplifies that both techniques demonstrated that there were good relationship existed between independent and dependent variables compare to RSM model which is r^2 (0.9820). The strength of the relationship is measured by the coefficient determination, therefore we can say that there was a strong positive correlation from both techniques.
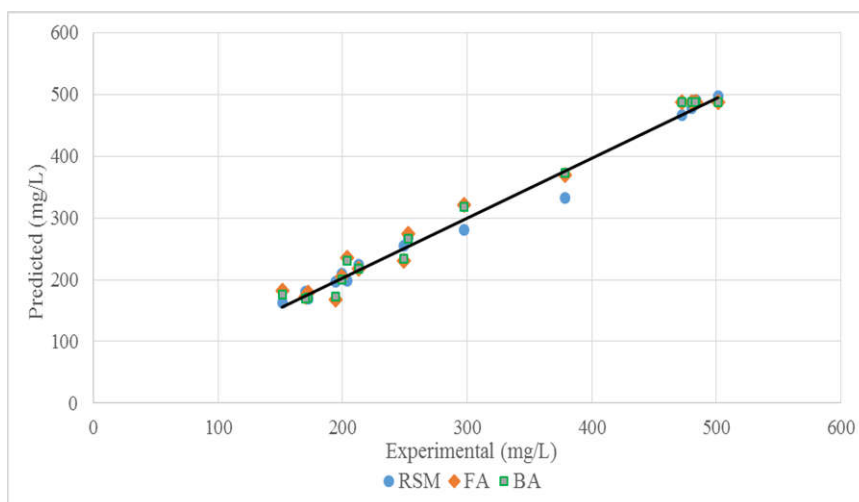
Fig. 7: Comparative plots of RSM, BA and FA

Fig. 7 shows the predicted flexirubin activity by RSM, FA and BA against the experimentally determined values. As observation in Fig. 7, most of the predicted results lie in the 45 degree line implying that the model predicted results are very close in agreement with the experimental data. It shows that BA and FA demonstrated that there were good relationship existed between independent and dependent variables compare to RSM model. Compared to metaheuristic techniques, BA was found to be highly significant than FA, and it indicator that FA generated has been slightly perfected to fit model, this is due to the fantastic behavior of echolocation of micro bats. The statistical analysis for RSM, FA and BA are shown in Table 5.

Table 5: Statistical analysis for RSM, FA and BA

Parameter	RSM	FA	BA
Correlation coefficient	0.9820	0.9838	0.9887
Accuracy	0.9612	0.9466	0.9574
Maximum error	45.11	32.83	26.8366
MAE	10.9559	13.9012	11.2541
RMSE	14.93583	16.9042	13.928

As shown in Table 5, the accuracy for RSM, FA and BA were 0.9612, 0.9466 and 0.9574, respectively. The analysis shows that RSM has slightly better accuracy than FA and BA. Although RSM shows higher accuracy than FA and BA, in fermentation area, coefficient, r^2 plays important roles in determined the superiority of the model. As demonstrated in Table 5, r^2 for BA was significantly better than RSM and FA. In this study, accuracy demonstrated the closeness of the measurement of FA and BA predicted data to the actual experimental data. The accuracy for RSM, FA and BA also can be shown in Fig. 8 below.
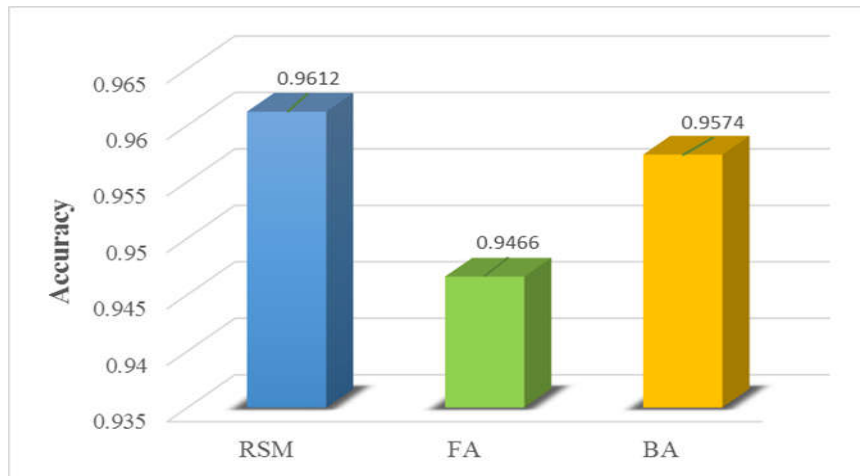
Fig. 8: Accuracy for RSM, FA and BA

For the validation data set, the maximum error for RSM, FA and BA were 45.11, 32.83 and 26.8366, respectively. It can be seen that the percent error for FA and BA were less than RSM, but BA shows better than FA. This exemplifies the capability of BA in predicting the accurate concentration of medium growth for optimum yield, compared to RSM and FA. Maximum error in this study indicated the highest error occur in the experiment. The maximum error for RSM, FA and BA are shown in Fig. 9.
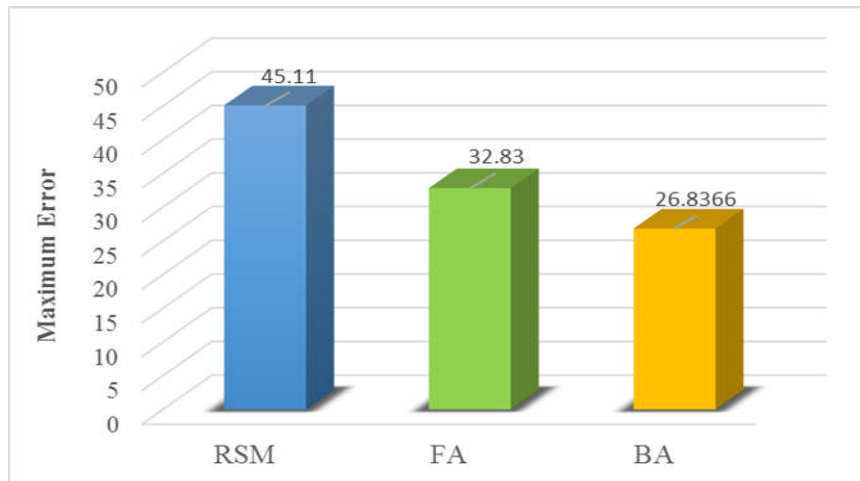

Fig. 9: Maximum error for RSM, FA and BA

From Table 5, mean absolute error (MAE) for BA is close to root mean square error (RMSE) (different 2.6739) which indicated that the BA based model only make small error, compared to RSM (different 3.9799) and FA (different 3.003).

Even with not much different between the predicted flexirubin by BA and FA, the results shows that BA based model has more precision in predicting the flexirubin than FA. Fig. 10 shows the MAE and RMSE for RSM, FA and BA techniques.

Fig. 10: MAE and RMSE for RSM, FA and BA

In conclusion, the wet-lab experiment and predicted data reveal that BA was an alternative technique for the best yield production. Indeed, FA and BA are alternative tools that provide quick optimization finding the optimize response variables on *flexirubin* production, compared to the traditional method (RSM).

5   Conclusion and Future Work

In this paper, three optimization techniques, RSM, FA and BA were applied for medium optimization in order to enhance the pigment (*flexirubin*) production by *C. artocarpi* CECT 8497[T]. The performance of the algorithms is evaluated by various statistical analyses and the authenticity have been proved by wet lab experiment. The results presented in this study indicate that BA provides a dynamic model for successful prediction of optimal operating conditions on yellowish-orange pigment production with a high correlation compared with well-known FA and RSM technique.

On the other hand, at present there are no reports in optimized *flexirubin* production using BA-based techniques. Consequently, this study is giving an insight into the *flexirubin* production area by utilizing the artificial intelligence based optimization techniques. Thus, it can be concluded that metaheuristic algorithm (BA) may present a better alternative for the bio-process community. However, the limitation of this paper is its only focus on FA and BA algorithms. Therefore, for future work perspectives we will conduct more experiments using others metaheuristics algorithm. We will also reduce the number of experiments

using a metaheuristics algorithm as well reduce the cost and increase the production of pigments.

References

[1] Ahmad, W. A., Yusof, N. Z., Nordin, N., Zakaria, Z. A., Rezali, M. F. (2012). Production and characterization of violacein by locally isolated Chromobacterium violaceum grown in agricultural wastes. *Applied biochemistry and biotechnology*, *167*(5), 1220-1234.

[2] Pal, M. P., Vaidya, B. K., Desai, K. M., Joshi, R. M., Nene, S. N., & Kulkarni, B. D. (2009). Media optimization for biosurfactant production by Rhodococcus erythropolis MTCC 2794: artificial intelligence versus a statistical approach. *Journal of industrial microbiology & biotechnology*, *36*(5), 747-756

[3] Desai, K. M., Survase, S. A., Saudagar, P. S., Lele, S. S., & Singhal, R. S. (2008). Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: case study of fermentative production of scleroglucan. *Biochemical Engineering Journal*, *41*(3), 266-273.

[4] Yang, X. S., He, X. (2013). Firefly algorithm: recent advances and applications. *International Journal of Swarm Intelligence*, *1*(1), 36-50.

[5] Roeva, O., & Slavov, T. (2012, September). Firefly Algorithm Tuning of PID Controller for Glucose Concentration Control during E. coli Fed-batch Cultivation Process. In *FedCSIS* (pp. 455-462).

[6] Zamani, M. K. M., Musirin, I., Hassan, H., Shaaya, S. A., Sulaiman, S. I., Ghani, N. A. M., & Suliman, S. I. (2018). Active and Reactive Power Scheduling Optimization using Firefly Algorithm to Improve Voltage Stability under Load Demand Variation. *Indonesian Journal of Electrical Engineering and Computer Science*, *9*(2), 365-372.

[7] Raja, S. B., Pramod, C. S., Krishna, K. V., Ragunathan, A., & Vinesh, S. (2015). Optimization of electrical discharge machining parameters on hardened die steel using Firefly Algorithm. *Engineering with Computers*, *31*(1), 1-9.

[8] Alim, A. M., Ponnambalam, S. G., & Kanagaraj, G. (2015, December). Robot Workcell Layout Optimization Using Firefly Algorithm. In *International*

Conference on Swarm, Evolutionary, and Memetic Computing (pp. 188-200). Springer, Cham.

[9] Yang, X. S., & He, X. (2013). Bat algorithm: literature review and applications. *International Journal of Bio-Inspired Computation*, *5*(3), 141-149.

[10] Ali, E. S. (2014). Optimization of power system stabilizers using BAT search algorithm. *International Journal of Electrical Power & Energy Systems*, *61*, 683-690.

[11] Dao, T. K., Pan, T. S., & Pan, J. S. (2018). Parallel bat algorithm for optimizing makespan in job shop scheduling problems. *Journal of Intelligent Manufacturing*, *29*(2), 451-462.

[12] Perez, J., Melin, P., Castillo, O., Valdez, F., Gonzalez, C., & Martinez, G. (2017, October). Trajectory Optimization for an Autonomous Mobile Robot Using the Bat Algorithm. In *North American Fuzzy Information Processing Society Annual Conference* (pp. 232-241). Springer, Cham.

[13] Abdel-Fattah, Y. R., Saeed, H. M., Gohar, Y. M., & El-Baz, M. A. (2005). Improved production of Pseudomonas aeruginosa uricase by optimization of process parameters through statistical experimental designs. *Process Biochemistry*, *40*(5), 1707-1714.

[14] Venil, C. K., Zakaria, Z. A., & Ahmad, W. A. (2015). Optimization of culture conditions for flexirubin production by Chryseobacterium artocarpi CECT 8497 using response surface methodology. *Acta Biochimica Polonica*, *62*(2).

[15] Venil, C. K., Zakaria, Z. A., Usha, R., & Ahmad, W. A. (2014). Isolation and characterization of flexirubin type pigment from Chryseobacterium sp. UTM-3T. *Biocatalysis and Agricultural Biotechnology*, *3*(4), 103-107.

[16] Singh, N., Goel, G., Singh, N., Pathak, B. K., & Kaushik, D. (2015). Modeling the red pigment production by Monascus purpureus MTCC 369 by Artificial Neural Network using rice water based medium. *Food Bioscience*, *11*, 17-22.

[17] Sinha, K., Chowdhury, S., Saha, P. D., & Datta, S. (2013). Modeling of microwave-assisted extraction of natural dye from seeds of Bixa orellana (Annatto) using response surface methodology (RSM) and artificial neural network (ANN). *Industrial Crops and Products*, *41*, 165-171.

[18] Joshi, C., & Singhal, R. S. (2016). Modelling and optimization of zeaxanthin production by Paracoccus zeaxanthinifaciens ATCC 21588 using hybrid genetic algorithm techniques. *Biocatalysis and Agricultural Biotechnology*, *8*, 228-235.

[19] Goel, N., Gupta, D., & Goel, S. (2013). Performance of firefly and bat algorithm for unconstrained optimization problems. *International Journal of Advanced Research in Computer Science and Software Engineering*, *31*, 1405-1409.