JOURNAL OF FUNDAMENTAL MATHEMATICS
AND APPLICATIONS (JFMA) VOL. 6 NO. 2 (NOV 2023)
Available online at www.jfma.math.fsm.undip.ac.id

FLOWER POLLINATION ALGORITHM (FPA): COMPARING SWITCH PROBABILITY BETWEEN CONSTANT 0.8 AND DOUBLE EXPONENT

Yuli Sri Afrianti[1*], Fadhil Hanif Sulaiman[2]

[1]*Statistics Scientific Group, Faculty of Mathematics and Natural Sciences, Bandung Institute of Technology*
[2]*Mathematics Undergraduate Program, Faculty of Mathematics and Natural Sciences, Bandung Institute of Technology*

Email: [1]yuli.afrianti@itb.ac.id , [2]fadhilhs@students.itb.ac.id
[*]Corresponding Author

Abstract. Flower Pollination Algorithm (FPA) is an optimization method that adopts the way flower pollination works by selecting switch probabilities to determine the global or local optimization process. The choice of switch probability value will influence the number of iterations required to reach the optimum value. In several previous literatures, the switch probability value was always chosen as 0.8 because naturally the global probability is greater than local. In this article, comparison is studied to determine the switch probability by using the Double Exponent rule. The results are analyzed using Hypothesis Testing to test whether there is a significant difference between the optimization results. The study involved ten testing functions, and results showed that the 0.8 treatment is significantly different from the Double Exponent. However, in general no treatment is better than the other.

Keywords: *Flower Pollination Algorithm, Switch Probability, Double Exponent, Hypothesis Testing, Inferential Statistics.*

I. INTRODUCTION

Optimization is a method to get the best or optimum result in the form of minimum or maximum values while not breaking any restrictions that may exist [1]. Optimization problems are often encountered even in cases related to daily activities. Mathematically, the optimization problem can be written as follows:

$$
\begin{array}{ll}
\text{minimum} & f_i(x), (i = 1,2,\ldots,I) \\
x \in \mathrm{R_d} & \Phi_j(x) = 0, (j = 1,2,\ldots,J) \\
\text{subject to} & \psi_k(x) \leq 0, (k = 1,2,\ldots,K)
\end{array}
$$

with $f_i(x)$, $\Phi_j(x)$, and $\psi_k(x)$ are functions of the decision vector $x = (x_1, x_2, \ldots, x_\mathrm{d})^\mathrm{T}$. Variable x_p with $\mathrm{p} = 1,2,\ldots,\mathrm{d}$ is called the decision variable of x. Function $f_i(x)$ is called the objective function, and $\mathrm{R_d}$ is the space spanned by x_p also called the decision space. Meanwhile, the space formed by the values of the objective function is called the solution space. Equation $\Phi_j(x)$ and inequalities $\psi_k(x)$ are called constraints. As a note, inequalities $\psi_k(x)$ above can also be constructed as ≥ 0, and objectives can be formulated as maximum.

Optimization problems can be solved using an optimization algorithm. An algorithm is a step-by-step procedure for providing calculations or instructions [2]. Since the discovery of heuristic algorithms by Alan Turing during World War II in an attempt to crack the German Enigma code, researchers have raced to develop such methods with the majority of them drawing inspiration from nature [3]. Some well-known metaheuristic algorithms include the Genetic Algorithm (GA) in 1960 which was based on Darwin's theory of evolution and natural selection, Particle swarm optimization (PSO) in 1995 which originated from the phenomenon of intelligence in flocks of fish and birds, and the Flower Pollination Algorithm (FPA) in 2012 which was inspired by the pollination process and characteristics of flowering plants [4], [5]. In finding the optimum value, heuristic and metaheuristic algorithms do not depend on the derivative of the objective function [6]. The optimum solution cannot be guaranteed, but a fairly good solution can be obtained.

The FPA algorithm was discovered by Xin-She Yang which generally has better performance than similar algorithms, namely Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). The concept of FPA is based on the flower pollination process and the characteristics of flowering plants. The process of flower pollination can be associated with the transfer of pollen [7]. In FPA, pollination is divided into two categories that is global and local pollination. The appearance of these two categories will be regulated by the switch probability (will be abbreviated as p-switch) which in the original research had a constant value $p = 0.8$. Furthermore, based on the FPA concept, the characteristic of flowering plants is flower constancy. Flower constancy is a phenomenon in which pollinators limit their visits to only a few plant species [8]. In FPA, the concept of flower constancy is defined as the probability of pollination that is proportional to the similarity between two flowers [9]. Although this algorithm has many advantages, the FPA method also has various disadvantages. Therefore, to improve its performance, the researchers made various modifications, some of which included combining it with similar algorithms such as GA and PSO, changing the step size, or changing the value of the p-switch [10]. Based on the background above, in this article we will compare p-switch in FPA that is constant 0.8 versus Double Exponent rule and the results will be analyzed using Hypothesis Testing to test whether there are significant differences between treatments due to p-switch modifications.

II. METHODOLOGY

FPA separates two categories of pollination, namely global and local. Global pollination contains biotic forms and methods of cross-pollination. This category is based on the long-distance nature of pollination. Mathematically, biotic pollinators such as bees will move according to the Lévy distribution, which is called Lévy flight. Meanwhile, local pollination contains abiotic forms and method of self-pollination. The distance of the pollination is nearby and can be modeled using a uniform distribution. Of course, these two categories cannot appear simultaneously, therefore the probability of the appearance of these two categories of pollination is regulated by a value called p-switch. Naturally, the value of local probabilities will be greater than global. The rules of the FPA are as follows:

1. Biotic and cross pollination are categorized as global pollination with the pollinator's movements following Lévy flight, that is, following the Lévy distribution.

2. Abiotic and self-pollination are categorized as local pollination which follows a uniform distribution.
3. Flower consistency can be interpreted as the probability of pollination being proportional to the similarity between two flowers.
4. Global and local pollination is regulated by the p-switch value, namely $p \in [0,1]$, with probability of local pollination is significantly greater.

In this article, the p-switch value in FPA will be compared. This comparation will be based on the special properties of the p-switch, that is probability value of local pollination will be greater than the global pollination [9]. The first p-switch value is 0.8 which is the recommendation from the original research. The second will be bassed on Khursheed et al. (2021) [10] and will be named as Double Exponent rule. Double Exponent considered produce faster convergence. In this modification, the global pollination probability is made such that its value decreases exponentially at each iteration. However, as a note, the p-switch value in this study is the probability of global pollination, which is the opposite of reference article. For this reason, changes are made which are expressed in the following equation:

$$R = p_{max} - \left(\frac{p_{max} - p_{min}}{N_{iter}}\right) * t, \tag{1}$$

$$p^{t+1} = 1 - exp(-exp(-R)), \tag{2}$$

with R is called performance index, N_{iter} is maximum iteration with recommended of 10000, t is number of iterations, p_{max} and p_{min} are value of maximum and minimum probability with value 1 and 0, and lastly p^{t+1} is a value of local probability on the $t+1$ iterations.

In this study, FPA will be simulated using various rules in the main reference [9]. There are three main rules to start the simulation. The rules are: simulation provisions, various objective functions (or just functions), and conditions of the objective functions (including: dimension, boundary, and minimum value). Table 1 explains the simulation provisions starting from optimization objectives to the number of simulations to be carried out. Next, let dimension denoted by variable d, then various objective functions to be simulated are presented in Table 2.

Table 1. Simulation provisions

No	Provision	Explanation
1	Optimization objective	Minimum.
2	Number of population (n)	25.
3	Error Tolerance (e)	10^{-5}.
4	Treatments	0.8 and Double Exponent
5	Maximum Iterations	200000.
6	Boundary Algorithm	Move the solution that is out on the boundary towards the boundary.
7	Mantegna Algorithm	An approach from the Lévy distribution to global pollination. A scale of 0.1 is used. [11]
8	Number of simulations	50 (with a specified RNG that is 1 to 50).

Table 2. Various objective functions to be tested

No	Name	Equation
1	Ackley	$f(\vec{x}) = -20exp\left[-0.2\sqrt{\frac{1}{d}\sum_{i=1}^{d} x_i^2}\right] - exp\left[\frac{1}{d}\sum_{i=1}^{d} cos(2\pi x_i)\right] + 20 + e$
2	Dejong	$f(\vec{x}) = \sum_{i=1}^{d} x_i^2$
3	Easom	$f(\vec{x}) = -cos(x_1)cos(x_2)exp[-(x_1-\pi)^2 - (x_2-\pi)^2]$
4	Griewangk	$f(\vec{x}) = \frac{1}{4000}\sum_{i=1}^{d} x_i^2 - \prod_{i=1}^{d} cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$
5	Michaelwicz	$f(\vec{x}) = -\sum_{i=1}^{d} sin(x_i) \cdot \left[sin\left(\frac{ix_i^2}{\pi}\right)\right]^{20}$
6	Rastrigin	$f(\vec{x}) = 10d + \sum_{i=1}^{d}[x_i^2 - 10cos(2\pi x_i)]$
7	Rosenbrock	$f(\vec{x}) = \sum_{i=1}^{d-1}[(x_i-1)^2 + 100(x_{i+1} - x_i^2)^2]$
8	Schwefel	$f(\vec{x}) = -\sum_{i=1}^{d} x_i sin\left(\sqrt{\lvert x_i \rvert}\right)$
9	Yang	$f(\vec{x}) = \left(\sum_{i=1}^{d} \lvert x_i \rvert\right) exp\left[-\sum_{i=1}^{d} sin(x_i^2)\right]$
10	Shubert	$f(\vec{x}) = \left[\sum_{k=1}^{5} kcos(k+(k+1)x_1)\right] \cdot \left[\sum_{k=1}^{5} kcos(k+(k+1)x_2)\right]$

In this simulation, various dimensions and different boundaries are used for each function. The boundaries can be defined as follows: suppose d is the dimension of a function f and $\vec{x} = (x_1, x_2, \ldots, x_d)$, then the boundary $[a, b]$ is for i = 1,2, …, d then $x_i \in [a, b]$. Conditions for each function can be seen in Table 3. It is necessary to know that the output of this simulation is the number of iterations to achieve optimum results. For this reason, the original FPA flowchart needs to be modified as seen in Figure 2.

Table 3. Conditions for each function

No	Name	Dimensions (d)	Boundary	Minimum Value (fmin)	Minimum Solution (xmin)
1	Ackley	128	[-5.12, 5.12]	0	(0, 0, …, 0)
2	Dejong	256	[-5.12, 5.12]	0	(0, 0, …, 0)
3	Easom	2	[-100, 100]	-1	(π, π)
4	Griewangk	2	[-600, 600]	0	(0, 0)
5	Michaelwicz	10	[0, π]	-9.6601517	(2.202906, 1.570796, 1.284992, 1.923058, 1.720470, 1.570796, 1.454414, 1.756087, 1.655717, 1.570796) [12]
6	Rastrigin	8	[-5.12, 5.12]	0	(0, 0)
7	Rosenbrock	16	[-5, 5]	0	(1, 1, …, 1)
8	Schwefel	128	[-500, 500]	-53629.8112	(420.9687, …, 420.9687)
9	Yang	16	[-2π, 2π]	0	(0, 0, …, 0)
10	Shubert	2	[-10, 10]	-186.7309	Has 18 global minimum

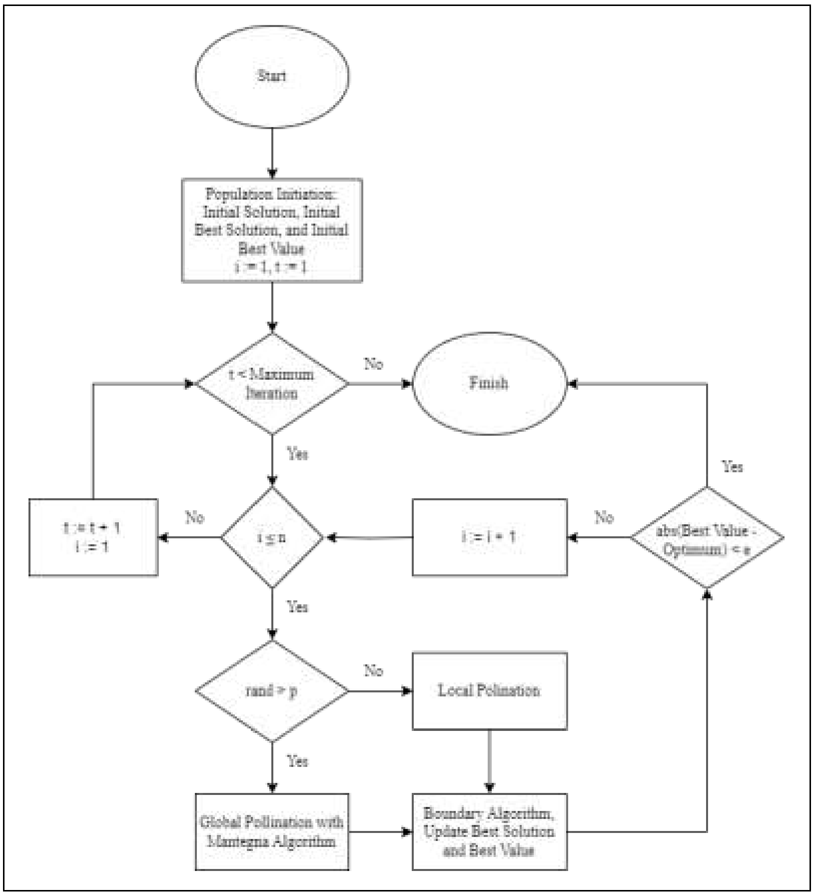

Figure 2. Modification of the FPA flowchart

Lastly the results of simulation will be analyzed using Hypothesis Testing [13], [14] to test whether there is a significant difference between treatments with p-switch 0.8 and Double Exponent (DE). The hypothesis test that will be used here is the t test with unknown but unequal variances. Each objective function will be tested by hypothesis:

$$H_0: \mu_{0.8} = \mu_{DE}$$
$$H_1: \mu_{0.8} \neq \mu_{DE}.$$

However, to be able to test this hypothesis, data should be normally distributed. Therefore, before testing the hypothesis, the data will be tested first using the Shapiro Wilk test [15]. The Shapiro Wilk test has a hypothesis:

$$H_0: The\ data\ comes\ from\ a\ normal\ distribution$$
$$H_1: The\ data\ does\ not\ comes\ from\ a\ normal\ distribution.$$

In the two tests above, a significance level of $\alpha = 5\%$ will be used.

III. RESULTS & DISCUSSION

The simulation results can be seen in Table 4. Note that, some preliminary results show that the success rate for achieving the optimal solution is still quite low. Even in the Rastrigin and Schwefel functions, the success rate was 0% for each treatment. This means that not all results obtained can be analyzed further. Therefore, the method will be carried out based on the minimum success rate criteria for all treatments that is 100%, 95% and 84%. To achieve this rate, dimension of every function that has success rate below 100% will be reduced one by one. After simulating with new dimensions and also cleaning out outliers, the results are presented in Table 4.

Tabel 4. Preliminary results written with: mean ± standard deviation (success rate%)

No	Name	0.8	Double Exponent
1	Ackley	12247.2 ± 996.4 (10%)	6233.6 ± 108.6 (100%)
2	Dejong	16867.9 ± 1039.7 (100%)	7843 ± 94 (100%)
3	Easom	195.1 ± 27.9 (100%)	254.3 ± 34.7 (100%)
4	Griewangk	443.9 ± 88.2 (100%)	603.4 ± 174 (100%)
5	Michaelwicz	NaN ± NaN (0%)	9622.5 ± 2682.1 (4%)
6	Rastrigin	NaN ± NaN (0%)	NaN ± NaN (0%)
7	Rosenbrock	3174.8 ± 336.2 (96%)	3635.4 ± 190.1 (100%)
8	Schwefel	NaN ± NaN (0%)	NaN ± NaN (0%)
9	Yang	94.1 ± 53.8 (100%)	448.1 ± 376.8 (100%)
10	Shubert	698.6 ± 167.1 (100%)	743.5 ± 471.3 (100%)

Table 5. Results for each criterion written with mean ± standard deviation (success rate%)

Criteria	No	Name (Dimension)	Treatment	
			0.8	Double Exponent
100%	1	Dejong (256)	16777.3 ± 827.8 (100%)	7843.4 ± 83.5 (100%)
	2	Easom (2)	196.8 ± 25.5 (100%)	254.9 ± 28.4 (100%)
	3	Griewangk (2)	443.9 ± 88.2 (100%)	603.4 ± 174 (100%)
	4	Yang (16)	83.8 ± 26.3 (100%)	344.6 ± 154.1 (100%)
	5	Shubert (2)	698.6 ± 167.1 (100%)	617.8 ± 284.8 (100%)
95%	1	Ackley (18)	1461.9 ± 101.4 (100%)	1180.4 ± 50.1 (100%)
	2	Michalewicz (5)	909 ± 133.5 (98%)	821.8 ± 162.9 (100%)
	3	Rastrigin (5)	1415 ± 212.4 (100%)	1592.2 ± 281.5 (100%)
	4	Rosenbrock (10)	1227.6 ± 95.2 (100%)	1863.4 ± 141 (100%)
84%	1	Ackley (24)	2093.8 ± 119.5 (98%)	1472.7 ± 40 (100%)
	2	Michalewicz (6)	1553.8 ± 288.1 (88%)	1755.6 ± 495.2 (100%)
	3	Rastrigin (6)	1949.7 ± 278.4 (92%)	2438.8 ± 429.7 (100%)
	4	Rosenbrock (33)	26593.5 ± 2959.6 (92%)	14449 ± 1529.7 (98%)

From Table 4, it can be seen that there are several results that are very different. For example, in the Dejong (256) function, the mean difference between the two treatments is around 8000 or in other words the 0.8 treatment is twice as slow as the Double Exponent. On the other hand, note that in the function Yang (16), it can be seen that the 0.8 treatment actually makes it three times faster than its opponent.

As has been discussed, these observations will be tested using hypothesis testing. However, before that, it is necessary to first test the normality of the data using the Shapiro Wilk test. The results of testing the normality assumption can be seen in Table 5. It can be seen that 26 data or 100% of the data are normally distributed at a significance level of 5%. This result is considered very good and meets the assumption of normality so it can be continued with Hypothesis Testing.

Table 6. P-value from the Shapiro Wilk Test

No	Name (Dimensions)	p-value	
		0.8	Double Exponent
1	Dejong (256)	1	1
2	Easom (2)	1	1
3	Griewangk (2)	0.282	0.065
4	Yang (16)	1	1
5	Shubert (2)	0.46	1
6	Ackley (18)	0.543	0.344
7	Michalewicz (5)	1	1
8	Rastrigin (5)	0.744	0.735
9	Rosenbrock (10)	1	1
10	Ackley (24)	1	1
11	Michalewicz (6)	1	0.355
12	Rastrigin (6)	1	0.26
13	Rosenbrock (33)	1	0.124

Table 7. P-value from t test for each function

Criteria	Nama Fungsi	$\bar{x}_{0.8}$	$\bar{x}_{DE}$	*p-value*	*p-value* (rounding)	*Conclusion* $(\alpha = 5\%)$
100%	Dejong (256)	16777.3	7843.4	5.12E-86	0.00	$\mu_{0.8} > \mu_{DE}$
	Easom (2)	196.8	254.9	9.31E-18	0.00	$\mu_{0.8} < \mu_{DE}$
	Griewangk (2)	443.9	603.4	8.87E-08	0.00	$\mu_{0.8} < \mu_{DE}$
	Yang (16)	83.8	344.6	3.21E-19	0.00	$\mu_{0.8} < \mu_{DE}$
	Shubert (2)	698.6	617.8	0.091185565	0.09	$\mu_{0.8} = \mu_{DE}$
95%	Ackley (18)	1461.9	1180.4	4.12E-32	0.00	$\mu_{0.8} > \mu_{DE}$
	Michalewicz (5)	909	821.8	0.006833377	0.01	$\mu_{0.8} > \mu_{DE}$
	Rastrigin (5)	1415	1592.2	0.000587711	0.00	$\mu_{0.8} < \mu_{DE}$
	Rosenbrock (10)	1227.6	1863.4	1.15E-43	0.00	$\mu_{0.8} < \mu_{DE}$
84%	Ackley (24)	2093.8	1472.7	2.60E-53	0.00	$\mu_{0.8} > \mu_{DE}$
	Michalewicz (6)	1553.8	1755.6	0.01974435	0.02	$\mu_{0.8} < \mu_{DE}$
	Rastrigin (6)	1949.7	2438.8	7.55E-09	0.00	$\mu_{0.8} < \mu_{DE}$
	Rosenbrock (33)	26593.5	14449	3.37E-43	0.00	$\mu_{0.8} > \mu_{DE}$

Observe Table 7, based on the p-value produced for each function, it can be seen that at the 5% significance level, almost all functions give results that the treatment means of 0.8 and Double Exponent are significantly different. However, only the Shubert function states that the treatment means of 0.8 is the same as Double Exponent. Therefore, it can be concluded that the 0.8 treatment is significantly different from the Double Exponent.

Tabel 8. The number of results for each hypothesis test conclusion with $\alpha = 5\%$ for each criterion

Criteria	**Conclusion**	**Banyakya hasil**
100%	$\mu_{0.8} < \mu_{DE}$	3
	$\mu_{0.8} = \mu_{DE}$	1
	$\mu_{0.8} > \mu_{DE}$	1
95%	$\mu_{0.8} < \mu_{DE}$	2
	$\mu_{0.8} = \mu_{DE}$	0
	$\mu_{0.8} > \mu_{DE}$	2
84%	$\mu_{0.8} < \mu_{DE}$	2
	$\mu_{0.8} = \mu_{DE}$	0
	$\mu_{0.8} > \mu_{DE}$	2

Tabel 9. The number of results for each hypothesis test conclusion with an $\alpha = 5\%$ in total

No	**Conclusion**	**Banyakya hasil**	**Percentage**
1	$\mu_{0.8} < \mu_{DE}$	7	53.85 %
2	$\mu_{0.8} = \mu_{DE}$	1	7.69 %
3	$\mu_{0.8} > \mu_{DE}$	5	38.46 %

Now we will see which treatment is better or in other words has a smaller mean. Tables 8 and 9 summarize how many of the results meet the conclusions of the hypothesis test. In table 8 which shows the number of conclusions for each criterion, note that at the 95% and 84% criteria the number of means that are larger or smaller is similar that is two function. However, at the 100% criterion the treatment of 0.8 is slightly better than with Double Exponent. Overall result that are summarized in Table 9, there are no big differences between treatments. In other words, even though the two treatments are very different, in general no treatment is better than the other.

IV. CONCLUSION

The comparison of p-switch between 0.8 by following the Double Exponent rule has been carried out. From the 10 functions tested in this article, it can be concluded that the 0.8 treatment is significantly different from the Double Exponent. However, even though the two treatments are very different, in general no treatment is better than the other.

REFERENCES

[1] A. M. Nassef, M. A. Abdelkareem, H. M. Maghrabie, and A. Baroutaji, “Review of Metaheuristic Optimization Algorithms for Power Systems Problems,” *Sustainability*, vol. 15, no. 12, p. 9434, Jun. 2023, doi: 10.3390/su15129434.

[2] X. Yang, "Nature-Inspired Optimization Algorithms (Elsevier Insights) (1st ed.)". *Elsevier*, 2014.

[3] X. Yang, *Engineering Optimization: An Introduction with Metaheuristic Applications*. 2010. [Online]. Available: http://ci.nii.ac.jp/ncid/BB03412264.

[4] X.-S. Yang, "Nature-inspired computation and Swarm Intelligence: Algorithms, theory and applications". *Academic Press*. 2020.

[5] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, M. A. Awadallah, and X. Yang, “Variants of the flower Pollination Algorithm: a review,” in *Studies in computational intelligence*, 2017, pp. 91–118. doi: 10.1007/978-3-319-67669-2_5.

[6] W. Wong and C. I. Ming, "A Review on Metaheuristic Algorithms: Recent Trends, Benchmarking and Applications," *2019 7th International Conference on Smart Computing & Communications (ICSCC)*, Sarawak, Malaysia, 2019, pp. 1-5, doi: 10.1109/ICSCC.2019.8843624.

[7] “Understanding flowers and flowering: an integrated approach,” *Choice Reviews Online*, vol. 52, no. 04, pp. 52–1978, Nov. 2014, doi: 10.5860/choice.185593.

[8] L. Chittka, J. D. Thomson, and N. M. Waser, “Flower constancy, insect psychology, and plant evolution,” *The Science of Nature*, vol. 86, no. 8, pp. 361–377, Aug. 1999, doi: 10.1007/s001140050636.

[9] X.-S. Yang, “Flower pollination algorithm for global optimization,” in *Lecture Notes in Computer Science*, 2012, pp. 240–249. doi: 10.1007/978-3-642-32894-7_27.

[10] M.-U.-N. Khursheed *et al.*, “PV Model parameter estimation using modified FPA with dynamic switch probability and step size function,” *IEEE Access*, vol. 9, pp. 42027–42044, Jan. 2021, doi: 10.1109/access.2021.3064757.

[11] R. N. Mantegna, “Fast, accurate algorithm for numerical simulation of Lévy stable stochastic processes,” *Physical Review*, vol. 49, no. 5, pp. 4677–4683, May 1994, doi: 10.1103/physreve.49.4677.

[12] C. Vanaret, “Certified Global MiNima for a benchmark of difficult optimization problems,” *arXiv.org*, Mar. 22, 2020. https://arxiv.org/abs/2003.09867.

[13] R. E. Walpole, *Probability and statistics for engineers and scientists*. 2012.

[14] Modul Praktikum Statistika KK Statistika, FMIPA ITB, 2014.

[15] S. S. Shapiro and M. B. Wilk, “An analysis of variance test for normality (complete samples),” *Biometrika*, vol. 52, no. 3–4, pp. 591–611, Dec. 1965, doi: 10.1093/biomet/52.3-4.591.