Flower Pollination Inspired Algorithm on Exchange Rates Prediction Case

I Nyoman Prayana Trisna*[1], Afiahayati[2], Muhammad Auzan[3]
[1]Information Technology Study Program, Udayana University, Bali, Indonesia
[2,3]Dept. of Computer Science and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia
e-mail: *[1]**prayana.trisna@unud.ac.id**, [2]afia@ugm.ac.id, [3]muhammadauzan@ugm.ac.id

Abstrak

Flower pollination algorithm (algoritma penyerbukan bunga) adalah sistem terinspirasi makhluk hidup yang memiliki proses yang mirip dengan algoritma genetika, yang bertujuan untuk memecahkan permasalahan optimisasi. Dalam penelitian ini, kami menguji pemanfaatan flower pollination algorithm dalam regresi linier untuk kasus pertukaran mata uang. Setiap solusi direpresentasikan dalam satu set koefisien regresi. Ukuran populasi untuk solusi dan switch probability antara global pollination (penyerbukan global) dan local pollination (penyerbukan local) diujicobakan dalam penelitian ini. Hasil penelitian ini menunjukkan bahwa solusi akhir diperoleh ketika menggunakan ukuran populasi tinggi dan switch probability yang tinggi. Selain itu, penelitian kami menemukan bahwa peningkatan ukuran populasi menyebabkan waktu eksekusi yang cukup lama, di mana kecendrungan untuk global pollination hanya sedikit meningkatkan waktu ekseksui.

***Kata kunci**—flower pollination algorithm, algoritma genetika, regresi linear, pertukaran mata uang, global pollination, local pollination*

Abstract

The flower pollination algorithm is a bio-inspired system that adapts a similar process to a genetic algorithm that aims for optimization problems. In this research, we examine the utilization of the flower pollination algorithm in linear regression for currency exchange cases. Each solution represents the regression coefficients. The population size for the solutions and the switching probability between global pollination and local pollination is experimented with in this research. The result shows that the final solution is obtained using a larger population and higher switch probability. Furthermore, our research finds that the increasing population size leads to considerable running time, where the probability of global pollination just slightly increases the running time.

***Keywords**—flower pollination algorithm, genetic algorithm, linear regression, exchange rate, global pollination, local pollination*

1. INTRODUCTION

The importance of planning a business target, resource plan, investing, and other management planning makes forecasting indispensable to ensure company continuity [1]. Even now, every significant company forecasts annually or quarterly and uses the computer science method to ensure better prediction. Computer science uses several methods to make a forecast model using several data, such as time-series data that even include an economic process in the calculation, with the easiest and lightest using only time-series data [2][3].

The forecasting model can be built using a traditional statistical model such as the autoregressive model [4][5], curve fitting [6], and nonlinear modeling such as neural networks [7][8]. Technology advancement in increasing computer speed and extensive data handling has given rise to another method also proposed in this research: a nature-inspired and intelligence-based model [9][10].

Much related research using the intelligence-based model have been published, such as comparing artificial neural network and hybrid ANN-ARMA and ANN-ARIMA [11][12]. The research results show that a hybrid method produces a model with less error than a pure ANN model. Adding ARIMA makes ANN robust, facing high volatility, complex, and noisy data [13]. Using a hybrid method, the model can predict with a correlation coefficient as significant as 99.545% [11].

Another method that has already been proven to forecast an exchange market value is the Genetic Algorithm Neural Network (GANN) [14]. GANN combines an intelligent Neural network and a nature-inspired method of the Genetic Algorithm [15]. Using one-year currency exchange data, the GANN can produce a model predicting the lowest RMSE of 0.044% [14].

In the emergence of state-of-the-art bio-inspired methodologies, Yang [16] proposed a new genetic algorithm method inspired by the behavior of flower pollination. Flower Pollination Algorithm is a swarm intelligence inspired by how flower pollination occurs. Flower pollination can occur by its flower or from another flower. In this algorithm, self-pollination is local search and pollination from another flower is global. The flower pollination algorithm solves the global and local search balance and uses levy for better global search performance. Yang [16] stated that there are four rules in the flower pollination algorithm:

- Rule 1 - Cross-pollination, or interfloral pollination, is a global pollination that follows a Levy distribution.

$$x_i^{t+1} = x_i^t + L(g^* - x_i^t) \quad (1)$$

$$L = \frac{\lambda \Gamma(\lambda) sin(\gamma \pi / 2)}{\pi s^{1+\lambda}} \quad (2)$$

- Rule 2 – Self-pollination is local pollination

$$x_i^{t+1} = x_i^t + \epsilon(x_j^t - x_k^t) \quad (3)$$

- Rule 3 – Pollinators can develop flower homeostasis because it corresponds to a reproductive probability proportional to the two flowers' similarity.

- Rule 4 – The interaction between self-pollination and global pollination can be controlled by random variables (which will be called switch probability), with slightly biased toward local pollination

In this research, the Flower Pollination Algorithm (FPA) model is deducted to forecast the exchange rate of the United Kingdom Pound sterling (GBP) to US dollars (USD). The exchange rate of GBP to USD in this research is a time series problem with several data before used as input, and the output is one step in the future.

A similar problem with exchange rates has been researched using classical methods like ARIMA [17] or hybrid methods by combining artificial intelligence with nature-inspired algorithms. An artificial neural network can be combined with classical statistical methods such as ARMA and ARIMA to achieve a robust forecast model [17][18]. The classical method is used to understand better the relationship between variables besides the ANN that works as an optimizer. However, using a hybrid method means increasing parameters and needing a new method to determine the parameter as in prior studies [19]. Furthermore, a hybrid of ANN with ARMA or ARIMA constructs a model that can withstand volatile, complex, and noisy data.

The genetic algorithm is a nature-inspired algorithm that uses the same principle as the classical or neural-based approach and can be used as a forecasting tool [21]. Like a flower pollination algorithm, a genetic algorithm is a population-based optimization technique initiated

from a random population and solution [22]. At each iteration happens reproduction wherein genetic algorithm only reproduces between the selected population to breed new and hopefully better solutions. In the genetic algorithm, the crossovers and mutations reproduce the next-generation solution from the prior solution [23]. However, in FPA, two crucial phases happen that is local pollination and global pollination [16]. Global pollination means pollination between flowers, and local pollination is pollination with itself. The mathematical equation of global and local pollination is mentioned in Equation 1 and Equation 3, respectively. Using Levy distribution [24] to perform diversification in global pollination or exploration makes the flower pollination algorithm more reliable in solving the global optimization problem. It guarantees the system to jump out of local optima [25].

　　　Time-series problems are nonlinear problems but with a relatively small dimension. Solving time-series problems using a neural-based approach or hybrid method will take too much effort compared to its dimension. Therefore, in this work, we propose only using FPA to find the optimal global solution to the time-series problem. Nonetheless, in this work, we proposed forecasting methods using flower pollination because it involves optimization like ANN and the survival of the fittest without forfeiting the computational cost.

2. METHODS

2.1 Dataset

　　　This research uses exchange rates of United Kingdom GBP to US USD. The exchange rates are from January 24th, 2008, until February 29th, 2016. From this date range, we obtain 2102 records in time series, which is presented in Figure 1.
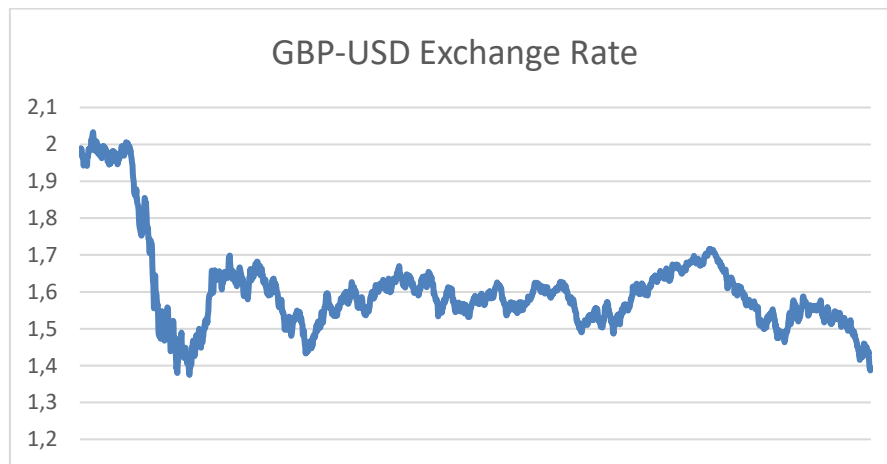

Figure 1 The GBP-USD data in the line chart

　　　For this research, we use the closing rates for each day. We conduct exploratory data analysis to get the statistical Information of the dataset. The result is shown in Table 1.

Table 1 Exploratory Data Analysis for the Dataset

Metric	Score
Mean	1.6057
Mode	1.5955
Median	1.58592
Standard Deviation	0.1229
Max	2.0331
Min	1.3749

From the exploratory data analysis, the peak exchange between GBP and USD is nearly twice, whereas the lowest exchange is when GBP is around 1.3 times USD. By measuring the mean, mode, median, and standard deviation of the dataset, the central tendency of the exchange rate in this dataset is around 1.6, with a relatively small deviation from the central tendency.

2.2 Flower Pollination Algorithm

The flower pollination algorithm is a bio-inspired system for optimization inspired by the characteristic of cross-pollination of flowers, which is proposed by Yang [16]. In his research, the flower pollination algorithm is written in pseudocode in Figure 2 and can be visualized in a flowchart as in prior research [22] as in Figure 3.

```
Flower Pollination Algorithm (or simply Flower Algorithm)
```
$Objective$ min or max $f(\mathbf{x})$, $\mathbf{x} = (x_1, x_2, ..., x_d)$
$Initialize$ a population of n flowers/pollen gametes with random solutions
$Find$ the best solution $\mathbf{g}_*$ in the initial population
$Define$ a switch probability $p \in [0, 1]$
while $(t < MaxGeneration)$
　　for $i = 1 : n$ (all n flowers in the population)
　　　　if $rand < p$,
　　　　　　$Draw$ a (d-dimensional) step vector L which obeys a Lévy distribution
　　　　　　$Global$ pollination via $\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + L(\mathbf{g}_* - \mathbf{x}_i^t)$
　　　　else
　　　　　　$Draw$ ϵ from a uniform distribution in [0,1]
　　　　　　$Randomly$ choose j and k among all the solutions
　　　　　　Do local pollination via $\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \epsilon(\mathbf{x}_j^t - \mathbf{x}_k^t)$
　　　　end if
　　　　$Evaluate$ new solutions
　　　　If new solutions are better, update them in the population
　　end for
　　$Find$ the current best solution $\mathbf{g}_*$
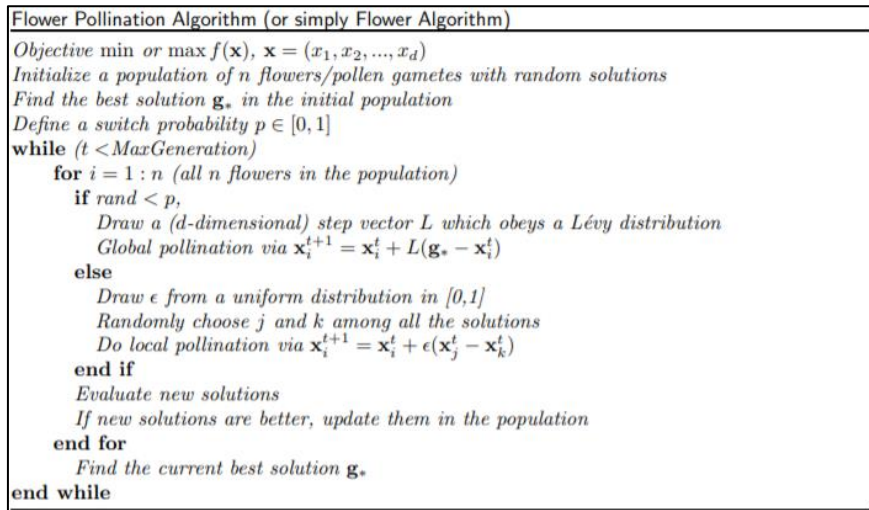**end while**

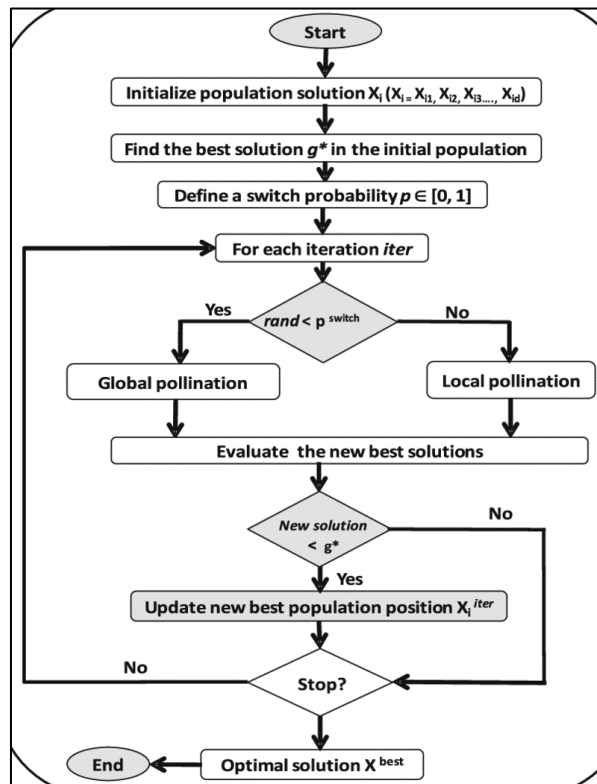Figure 2 Pseudocode of flower pollination algorithm by Yang [16]

Figure 3 Flowchart of flower pollination algorithm [22]

Our research aims to obtain the best solution $g*$ from the set of solutions x at generation t. From Figure 2, we initiate the initial population t_0 by generating d random solutions x, which will be explained in the next section. For every generation with d solutions, each solution x consists of regression coefficients x_i where $i = 1, 2, .., n$ and bias x_0 for predicting the exchange rate of day D'_T using linear regression based on n days before so that $x = \{x_0, x_1, x_2, ..., x_n\}$. The x_i will be used as a sum product with D_{T-i}, and x_0 will be added. Formally, the linear regression for our research is written in Equation 4.

$$D'_T = x_0 + \sum_{i=1}^{n} x_j D_{T-i} \qquad (4)$$

The objective function for each solution x is to minimize the difference between the predicted exchange rate D'_T and the actual exchange rate D_T. For this research, we use rooted mean squared error (RMSE) to measure the difference for each solution. RMSE is written in Equation 5, where n is equal to the length of the time series record. In the ideal result, the RMSE from Equation 5 will result in 0 because the same value between the predicted exchange rate D'_T and actual exchange rate D_T, so the squared error became zero.

$$RMSE(x) = \sqrt{\sum_{i=1}^{n} \frac{\left(D'_i(x) - D_i\right)^2}{n}} \qquad (5)$$

By minimizing the RMSE for each solution, we state the fitness function for each solution as in Equation 6 to evaluate the solution. The best solution for every generation will be the $g*$ and will be used as the final solution. According to Equation 5, the RMSE will yield 0 in the ideal situation, thus resulting ideal fitness of 1.

$$fitness(x) = \frac{1}{1 + RMSE(x)} \qquad (6)$$

For each generation, we propose d solutions as a population. From the initial generation ($t = 0$), the best solution in the population will be stated as $g*$. In the generation t where $t = 1, 2, ..., MaxGeneration$, if one solution is better than $g*$, that solution will replace the new $g*$. The alteration of $g*$ is done iteratively in each generation, so a dynamic approach is needed. The solutions in generation t are formed from the pollination of the solutions in generation t-1, either global or local, as stated in Equations 1 and 2, respectively. The switch between global or local pollination in generation t is controlled by switch probability p. In the next section, we will explore the hyperparameters for this research.

2.3 Experiment Setup

The data is split into training data to obtain the most optimal coefficient as the solution and testing data to evaluate the final solution. The training and testing data comparison is 80:20. The data is split after we process the data into chunks of sliding windows with a size of n.

In this research, we experiment with some hyperparameters for the algorithm. We also state the fixed hyperparameters that do not undergo the experiment, such as the number of days used for predicting (n), which is 5, and the number of generations/iterations m equals 100. Population size d and switch probabilities p are hyperparameters that are fine-tuned. In this research, we experiment with 50, 100, and 150 as the size of population d and 0.3, 0.5, and 0.8 as the switch probability p. From the combinations of d and p, and we obtain 9 experiment results. For each experiment, we measure the performance of the final solution by measuring the RMSE and the running time.

3. RESULT AND DISCUSSION

The flower pollination algorithm gives fast convergence results when training to get the optimal solution. The fast convergence result is shown in Figure 4, which shows only to generation 50th. Even if we observe each experiment with 100 as m or the number of generations/iterations, all experiments have the same corresponding $g*$ from generation 44th to 100th, and this left a question of whether the flower pollination algorithm gives premature convergence. Furthermore, most experiments show that convergence happens before the 20th generation.
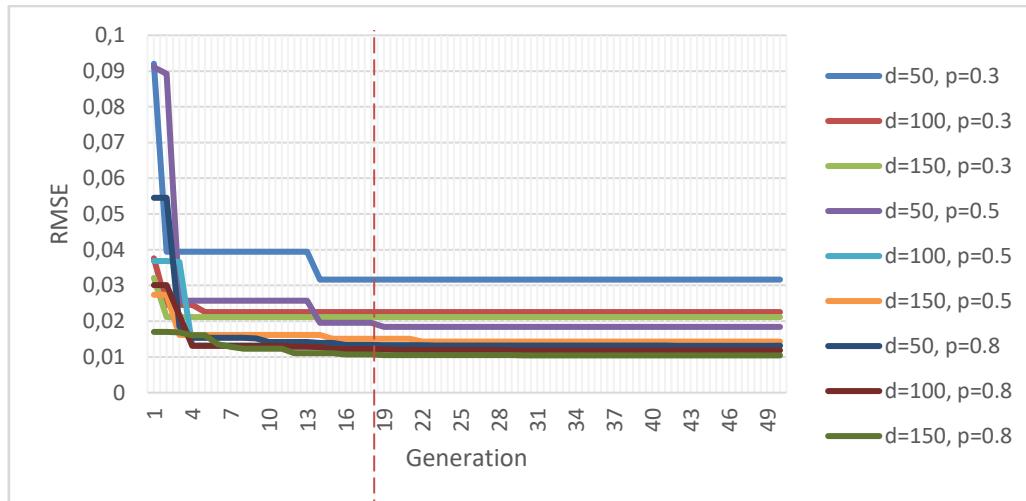
Figure 4 RMSE of $g*$ in each generation for each experiment of population d and switch probability p

Detailed results of RMSE in the final generation of testing data for all experiments are shown in Table 2. From this result, we can see that the population size is relatively inversely proportional to RMSE. Although the experiment with 0.5 as switch probability does not show the best RMSE when the population size is 150, the difference is insignificant when the population size equals 100. That means the population size is proportional to the solution's fitness, as we can see that 150 as the population size gives better results than 100 and 50 as the population size, while the population size at 100 is better than 50. This experiment shows that the number of formed solutions affects the best solution g*quality because the solutions in generation t are directly generated from pollination in generation t-1. The more solutions in generation t-1 mean more pollinations occur for generation t. More pollinations imply more probability of generating a better solution from prior $g*$.

Table 2 RMSE in the final generation of testing data

Population Size	Switch Probability	RMSE	Running time
50	0.3	0.034096	19.6428
50	0.5	0.018437	20.3304
50	0.8	0.013379	21.2921
100	0.3	0.023125	39.2658
100	0.5	0.013555	40.8589
100	0.8	0.011408	42.9194
150	0.3	0.021377	58.6585
150	0.5	0.015056	60.5198
150	0.8	0.010777	65.5033

For the switch probability *p*, we can see that a higher switch probability gives a better solution based on RMSE in Table 2. In each size of population *d*, higher *p* always gives a lower RMSE, which means a better solution. A higher *p* equals a higher possibility of global pollination to happens. That means global pollination tends to generate better solutions rather than local pollination. As stated in Equation 1, each solution of generation *t* is generated from pollinating the best solution *g** with itself at *t*-1. This solution tends to result in a solution close to g*, if not better than the previous *g**. Compared to global pollination, local pollination only pollinates random solutions to become a solution in the next generation. Even though it could generate a better solution than *g**, that probability is lower than global pollination, as our switch probability experiment proved.

Further analysis also can be conducted based on the prediction visualization. Figure 5, Figure 6, and Figure 7 show the prediction for each experiment. The final solution of each experiment is applied to all data. We take an average of 20 consecutive data in these visualizations to maintain the image quality.
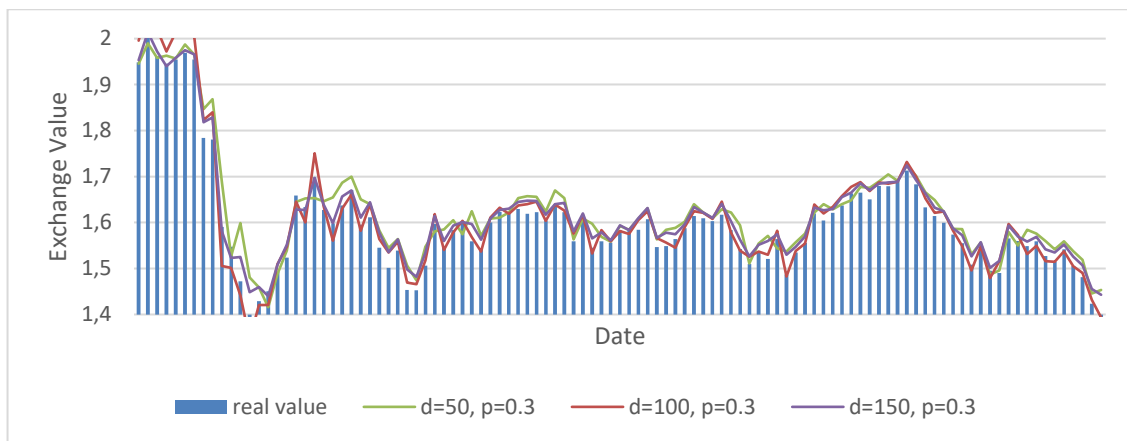

Figure 5 The prediction for each experiment of population *d* and switch probability *p* = 0.3
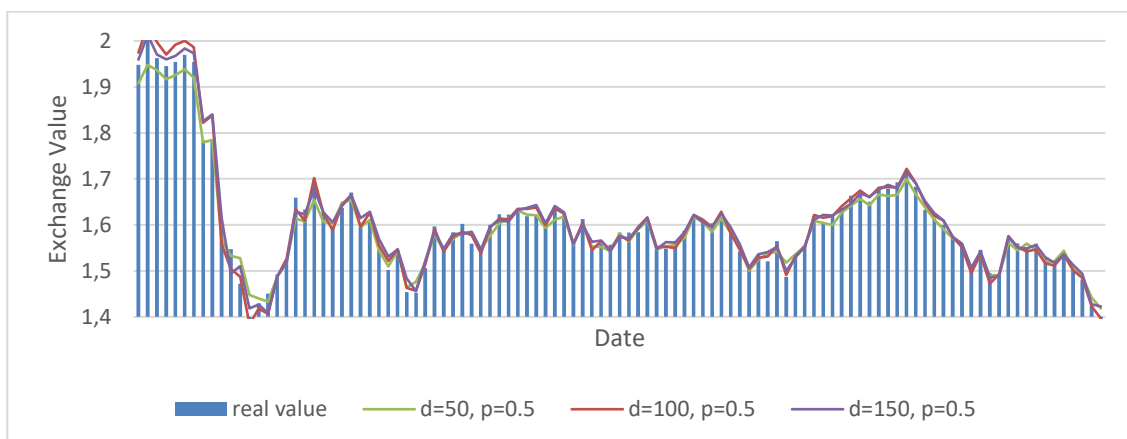

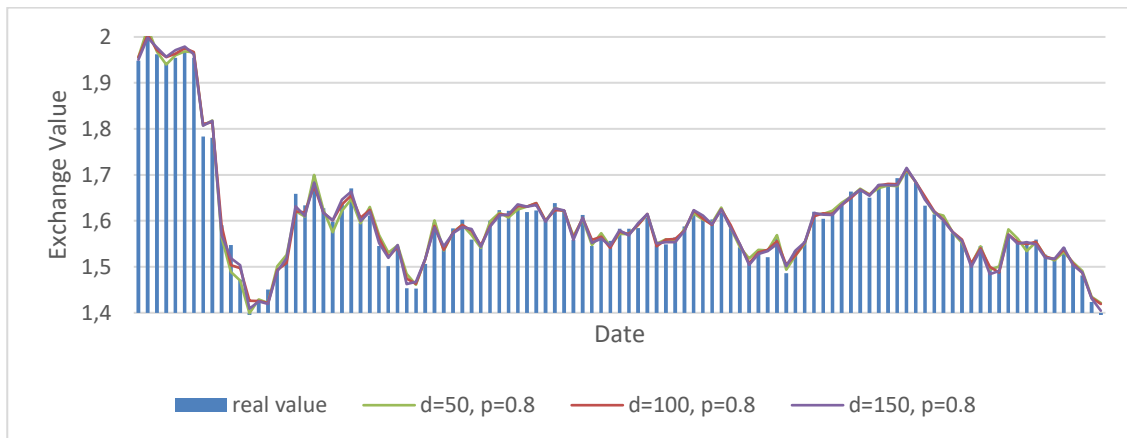Figure 6 The prediction for each experiment of population *d* and switch probability *p* = 0.5

Figure 7 The prediction for each experiment of population *d* and switch probability *p* = 0.8

From a brief view, the prediction of each experiment is similar to the actual data for each time in each experiment. However, it is noticeable that in Figure 5 (where *p* = 0.3), the prediction misses by a large margin when the actual value has sudden peaks or sudden drops from the previous actual value. The RMSE also reflects this from Table 2, where *p* = 0.3 gives the worst RMSE. The solution in Figure 6 improves the miss-prediction in Figure 5 by increasing the switching probability *p* to 0.5, and further improvement is shown in Figure 7, which uses *p* = 0.8. The solution is also reflected in Table 2, where *p* = 0.8 returns the best result. The process yields another question of whether using entirely global pollination for the flower pollination algorithm is better in this exchange rates case.

The training time for all experiments is also recorded and written in Table 2. From Table 2, running time is proportional to the switch probability *p*. Running time is longer when *p* is more significant in equal population size, and vice versa, indicating that global pollination consumes slightly more computational time. However, the running time difference of each experiment based on switch probability is insignificant, as shown in Table 2. From this experiment, it can be concluded that global pollination and local pollination serve relatively similar –if not the same time complexity. The noticeable time difference between population size d can be seen with the constant switch probability. The running time is linearly proportional to the population size *d* since the greater population size means more time to evaluate each solution in a generation.

4. CONCLUSION

The flower pollination algorithm is an algorithm for solving optimization problems. Cross-pollinating characteristics of flowers inspire that. Our research examines using a flower pollination algorithm for time series prediction in linear regression, where the solutions are represented in flowers. Each solution is a set of coefficients for regression, where we evaluate each flower in one generation to get the best solution using RMSE. The next generation is generated by cross-pollinating all solutions in one prior generation, either with global or local pollination. The best solution for all generations is considered the final solution.

This research conducts several experiments to examine the impact of population size and switch probability. All the experiments go with fast convergence, even below the maximum iteration. Our result shows that a higher population size gives the final solution with lower RMSE, which leads to higher fitness. The same conclusion goes to switch probabilities, where higher probabilities give better results in the final solution due to global pollination being likely to generate better solutions than the current best solution, rather than local pollination, which

pollinates the solution to other random solutions. In running time, increasing both switch probability and population size contributes to the increment of training time, although the population size leads to more running time than switch probability.

In the future, more experiments with flower pollination algorithms are needed. Further experiments can modify the global pollination that usually applies to Levy distribution. A similar attempt also can be made in local pollination by modifying the randomness of a picked solution to be cross-pollinated. Future research can also use the vast majority –if not entirely global pollination.

ACKNOWLEDGMENTS

REFERENCES

[1] T. Henttu-Aho, 'The role of rolling forecasting in budgetary control systems: Reactive and proactive types of planning', *Journal of management control*, vol. 29, no. 3–4, pp. 327–360, 2018.

[2] C. Foroni, M. Marcellino, and D. Stevanovic, 'Forecasting the Covid-19 recession and recovery: Lessons from the financial crisis', *International Journal of Forecasting*, vol. 38, no. 2, pp. 596–612, 2022.

[3] M. S. Islam and E. Hossain, 'Foreign exchange currency rate prediction using a GRU-LSTM hybrid network', *Soft Computing Letters*, vol. 3, p. 100009, 2021.

[4] M. Cai, M. Pipattanasomporn, and S. Rahman, 'Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques', *Applied energy*, vol. 236, pp. 1078–1088, 2019.

[5] B. Njindan Iyke, 'The disease outbreak channel of exchange rate return predictability: Evidence from COVID-19', *Emerging Markets Finance and Trade*, vol. 56, no. 10, pp. 2277–2297, 2020.

[6] S. K. Tamang, P. D. Singh, and B. Datta, 'Forecasting of Covid-19 cases based on prediction using artificial neural network curve fitting technique', *Global Journal of Environmental Science and Management*, vol. 6, no. Special(Covid-19), pp. 53–64, 2020.

[7] W. Cao, W. Zhu, W. Wang, Y. Demazeau, and C. Zhang, 'A deep coupled LSTM approach for USD/CNY exchange rate forecasting', *IEEE Intelligent Systems*, vol. 35, no. 2, pp. 43–53, 2020.

[8] P. L. Seabe, C. R. B. Moutsinga, and E. Pindza, 'Forecasting cryptocurrency prices using LSTM, GRU, and bi-directional LSTM: A deep learning approach', *Fractal and Fractional*, vol. 7, no. 2, p. 203, 2023.

[9] H. Lu, M. Azimi, and T. Iseley, 'Short-term load forecasting of urban gas using a hybrid model based on improved fruit fly optimization algorithm and support vector machine', *Energy Reports*, vol. 5, pp. 666–677, 2019.

[10] N. Nabipour, M. Dehghani, A. Mosavi, and S. Shamshirband, 'Short-term hydrological drought forecasting based on different nature-inspired optimization algorithms hybridized with artificial neural networks', *IEEE Access*, vol. 8, pp. 15210–15222, 2020.

[11] S. Wang, Z. Tang, and B. Chai, 'Exchange rate prediction model analysis based on improved artificial neural network algorithm', in *2016 International Conference on*

Communication and Electronics Systems (ICCES), 2016, pp. 1–5.

[12] S. A. Shahriar *et al.*, 'Potential of Arima-ann, Arima-SVM, dt and catboost for atmospheric PM2. 5 forecasting in bangladesh', *Atmosphere*, vol. 12, no. 1, p. 100, 2021.

[13] M. Durairaj and B. H. K. Mohan, 'A review of two decades of deep learning hybrids for financial time series prediction', *International Journal on Emerging Technologies*, vol. 10, no. 3, pp. 324–331, 2019.

[14] A. Sespajayadi, I. Nurtanio, and Others, 'Technical data analysis for movement prediction of Euro to USD using genetic algorithm-neural network', in *2015 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 2015, pp. 23–26.

[15] D. K. Sharma, H. S. Hota, K. Brown, and R. Handa, 'Integration of genetic algorithm with artificial neural network for stock market forecasting', *International Journal of System Assurance Engineering and Management*, vol. 13, no. Suppl 2, pp. 828–841, 2022.

[16] X.-S. Yang, 'Flower pollination algorithm for global optimization', in *Unconventional Computation and Natural Computation: 11th International Conference, UCNC 2012, Orléan, France, September 3-7, 2012. Proceedings 11*, 2012, pp. 240–249.

[17] A. Qonita, A. G. Pertiwi, and T. Widiyaningtyas, 'Prediction of rupiah against us dollar by using arima', in *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2017, pp. 1–5.

[18] F. Raheem and N. Iqbal, 'Forecasting foreign exchange rate: Use of FbProphet', in *2021 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, 2021, vol. 4, pp. 44–48.

[19] R. S. Al-Gounmeein and M. T. Ismail, 'Forecasting the exchange rate of the Jordanian Dinar versus the US dollar using a Box-Jenkins seasonal ARIMA model', *International Journal of Mathematics and Computer Science*, vol. 15, no. 1, pp. 27–40, 2020.

[20] E. Dave, A. Leonardo, M. Jeanice, and N. Hanafiah, 'Forecasting Indonesia exports using a hybrid model ARIMA-LSTM', *Procedia Computer Science*, vol. 179, pp. 480–487, 2021.

[21] M. Zivkovic, N. Bacanin, A. Djordjevic, M. Antonijevic, I. Strumberger, and T. A. Rashid, 'Hybrid genetic algorithm and machine learning method for covid-19 cases prediction', in *Proceedings of international conference on sustainable expert systems: ICSES 2020*, 2021, pp. 169–184.

[22] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, M. A. Awadallah, and X.-S. Yang, 'Variants of the flower pollination algorithm: a review', *Nature-inspired algorithms and applied optimization*, pp. 91–118, 2018.

[23] A. Hassanat, K. Almohammadi, E. Alkafaween, E. Abunawas, A. Hammouri, and V. B. S. Prasath, 'Choosing mutation and crossover ratios for genetic algorithms—a review with a new dynamic approach', *Information*, vol. 10, no. 12, p. 390, 2019.

[24] C. Tsallis, S. V. F. Levy, A. M. C. Souza, and R. Maynard, 'Statistical-mechanical foundation of the ubiquity of Lévy distributions in nature', *Physical Review Letters*, vol. 75, no. 20, p. 3589, 1995.

[25] M. Abdel-Baset and I. M Hezam, 'An effective hybrid flower pollination and genetic algorithm for constrained optimization problems', *Advanced Engineering Technology and Application*, vol. 4, no. 3, pp. 28–35, 2015.