

PAPER • OPEN ACCESS

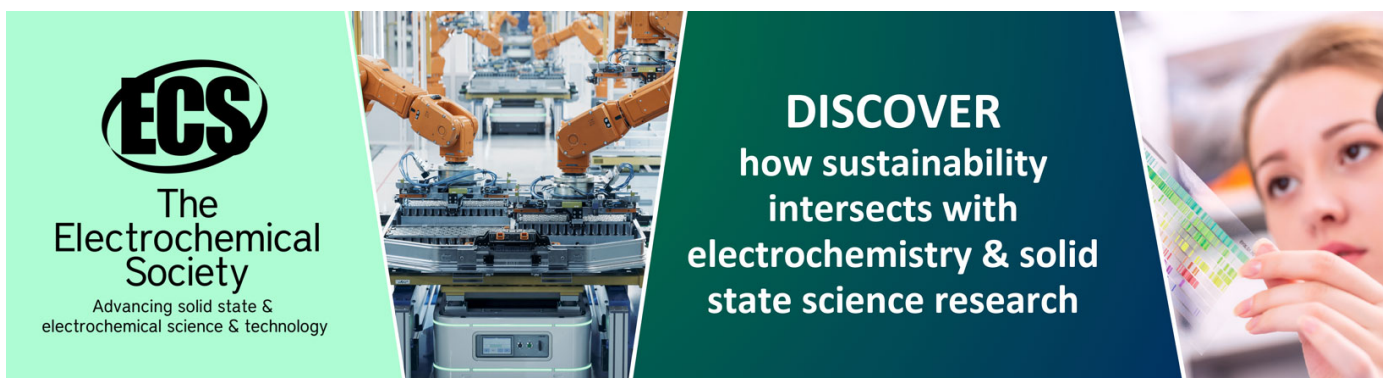
## Forecasting foreign exchange rate using a combination of linear regression and flower pollination algorithm

To cite this article: I B N Pascima and I M Putrama 2021 *J. Phys.: Conf. Ser.* **1810** 012021

View the [article online](#) for updates and enhancements.

### You may also like

- [Quantum versus classical generative modelling in finance](#)  
Brian Coyle, Maxwell Henderson, Justin Chan Jin Le et al.
- [An Improved Self-Adaptive Flower Pollination-Backpropagation Neural Network for the State of Charge Estimation of Lithium-Ion Batteries with Output Sliding Average Window Strategy](#)  
Yuanru Zou, Shunli Wang, Nan Hai et al.
- [International Currency Translator using IoT for shopping](#)  
CH.M.H Sai baba, Dr.S Hrushikesava Raju, M.V.B.T Santhi et al.



**ECS**  
The  
Electrochemical  
Society  
Advancing solid state &  
electrochemical science & technology

**DISCOVER**  
how sustainability  
intersects with  
electrochemistry & solid  
state science research

# Forecasting foreign exchange rate using a combination of linear regression and flower pollination algorithm

I B N Pascima<sup>1</sup> and I M Putrama<sup>2</sup>

<sup>1,2</sup>Faculty of Technology and Vocational, Universitas Pendidikan Ganesha, Bali, Indonesia, 81116

Email: [gus.pascima@undiksha.ac.id](mailto:gus.pascima@undiksha.ac.id), [made.putrama@undiksha.ac.id](mailto:made.putrama@undiksha.ac.id)

**Abstract.** Several currencies exist in the world. Each currency will have value. The currency exchange will go through a conversion process to adjust the amount. Each currency value can fluctuate based on the conditions of the currency area. The fluctuating changes in value provide profit opportunities. Maximizing profits can make forecasts so that the right decisions are made. One of the forecasts can use regression. Regression is capable of forecasting based on historical data. The regression in this study will be optimized using the Flower Pollination Algorithm (FPA). The use of the Flower Pollination Algorithm (FPA) aims to obtain appropriate parameters for regression to reduce forecast errors. The data in this study were obtained by utilizing extraction from the Meta trader application. This data will be the basis for the system learning stage and the testing phase. Obtaining a good hyperparameter can make the forecasting system closer to the actual value. Good system accuracy can be a trader's supporting data in making transactions. Forecasting in this study used the parameter 5 window sizes, 20 population sizes, and 0.7 probability switch. This experiment resulted in MSE 0.0331 and RMSE 0.1756. This forecasting has sufficient results to support a trader's decision. Further research is needed to improve accuracy and determine the direction of the forecast to improve this research.

## 1. Introduction

Money is one of the agreed instruments of exchange. In the world, there are various types of currencies. Each region will have its legal currency. Each currency has its value. If a currency is exchanged for another currency, a conversion will occur. Currency exchange rates will be relative to other currencies and the value fluctuates. This change is influenced by many things such as the economic condition of a country, political conditions, natural disasters that occur in an area, etc. The difference in exchange rates is what triggers the Foreign Exchange (Forex).

The value in foreign currency fluctuates so much that there will be price differences that can be used to get profit. It is the same as selling gold, which can have a difference in value when buying and selling it. A similar thing is experienced in stock trading, where prices can change at any time. The advantage obtained from the difference in foreign currency value is what attracts someone to buy and sell currencies. A person who trades currency is often referred to as a trader.

Traders carry out their activities by utilizing information, analysis, and intuition to make a profit. Traders use and analyze this information often termed fundamental analysis. Apart from fundamental analysis, it is also known as technical analysis. Technical analysis is a mechanism to compare currency exchange rates based on historical data or previous currency exchange rates. Technical



analysis is often performed to provide confidence in a trader's decision. Every trader will carry out various analyzes to get the right decision. This is useful for increasing profits or reducing the possibility of losses.

In the foreign exchange market, technology has made it easier for transactions to occur. The use of technology also makes transactions happen faster. Currency exchange rate fluctuations can be monitored in seconds. Transaction speed can provide high profit or loss opportunities. Seeing the risks that exist in foreign exchange, a good analysis is needed to deal with risks. Apart from fundamental analysis, technical analysis is also often used to support decisions. There are various tools to support technical analysis. Analysis can be done with the help of tools and mathematical calculations, while fundamental analysis will depend on a trader's ability to read the market.

Technical analysis uses historical data for forecasting. Traders tend to find it difficult to make forecasts without forecasting tools or methods. Based on the foregoing, the writer tries to develop a forecasting optimization using one of the methods in the Artificial Intelligence (AI) branch. The method used is the Flower Pollination Algorithm (FPA). Flower Pollination Algorithm (FPA) was developed by Xin-She Yang in 2012 [1]. FPA is a method that is inspired by living things. This algorithm works by simulating the pollination process of flowers. FPA works very efficiently with the use of a few iterations with good results [2]. The two main steps in FPA are global pollination and local pollination. In the global pollination step, pollen is carried by pollinators and can travel long distances. This ensures optimal pollination and reproduction (the best fitness value) and the fitness value is represented as  $g^*$  [3].  $g^*$  is the best flower for global solutions.

This method is not a forecasting method but rather an optimization method. Based on the characteristics of the method, this method is estimated to be able to perform optimization on other forecasting methods. So, this method will be combined with other forecasting methods. The forecasting method used is a regression which will be optimized using FPA. Similar research has been conducted by [4] using Genetic Algorithms. This combination is expected to be able to obtain a small forecast error value to minimize the possibility of trader losses and maximize profits. Through this research, it is expected to be able to describe the ability of FPA optimization for time series data characters in the linear regression method. In addition to describing optimization capabilities, it is hoped that a good forecasting model for forex data is also expected to be able to help traders make decisions.

## 2. Research Method

Figure 2 is a forecast model that combines linear regression with the Flower Pollination Algorithm. The system workflow begins with data collection obtained from the Metatrader application which is used as training data to obtain the best hyperparameters. After the data are obtained, the normalization process is carried out using Equation (1) to the data. After the normalization process, it is continued to check the hyperparameters that are owned. If there are no hyperparameters, then proceed with filtering the data and dividing the data into training data and validation data. The system will carry out the training process using training data and perform a validation process where the hyperparameters will be saved in \*.txt format. If the hyperparameter matches produce tolerable errors, then proceed to the forecasting stage using data testing. The value of the forecast results will be converted back to the initial value through the denormalization process using Equation (2).

### 2.1. Data

The development of this system is based on a technical analysis approach by utilizing historical data. This system uses optimized regression using FPA in determining predictions. The historical data used in this study were obtained from the Metatrader application. Data taken with the Metatrader application were daily EURUSD data for 3 years from January 2017 to December 2019. The data variables used in this study only uses the daily forex closing value. The data display is shown in Figure 1.

| Time       | Open    | High    | Low     | Close   | Volume |
|------------|---------|---------|---------|---------|--------|
| 2017.01.02 | 1.051   | 1.05275 | 1.0448  | 1.04519 | 72718  |
| 2017.01.03 | 1.04524 | 1.04898 | 1.03399 | 1.04044 | 139324 |
| 2017.01.04 | 1.04044 | 1.04996 | 1.03895 | 1.04871 | 138590 |
| 2017.01.05 | 1.04871 | 1.06149 | 1.04804 | 1.06055 | 170775 |
| 2017.01.06 | 1.06005 | 1.06207 | 1.05235 | 1.05235 | 146005 |
| 2017.01.09 | 1.05292 | 1.05826 | 1.05104 | 1.05723 | 125786 |
| 2017.01.10 | 1.05718 | 1.06268 | 1.05506 | 1.05534 | 147521 |

**Figure 1.** Data daily EURUSD.

## 2.2. Normalization

All data, both training and validation data, will go through the normalization process. This process aims to convert the original data to the same range as the range used. There are various normalization methods, one of which gives the best accuracy is the Min-max method which changes to a range of [0, 1] [5]. Normalization uses the formula in Equation 1 [6].

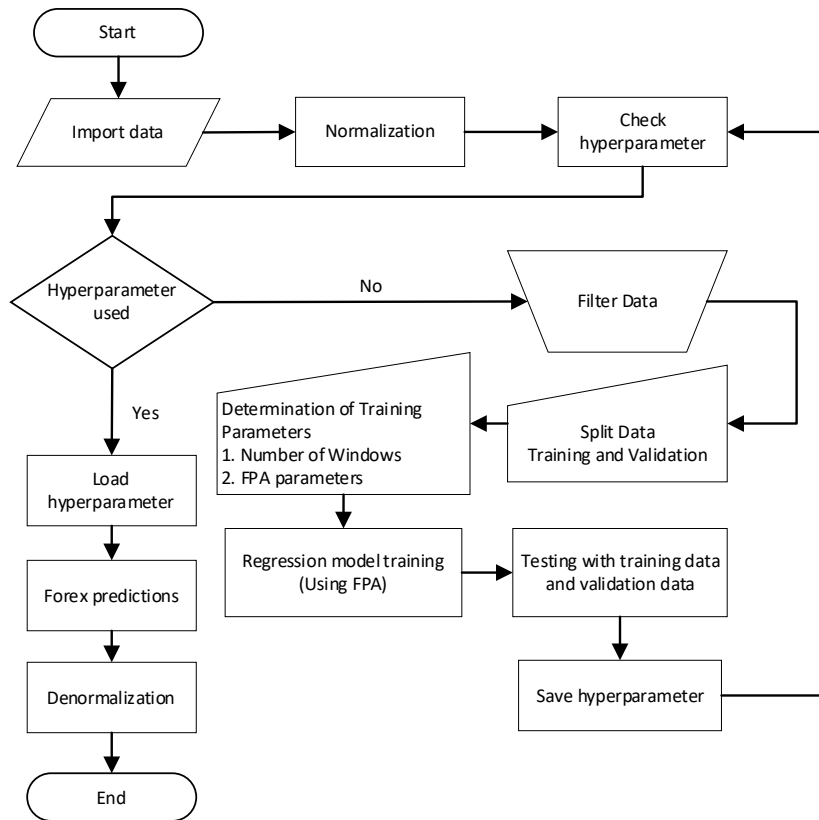
$$x' = \left( \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \right) * (\max - \min) + \min \quad (1)$$

where  $x'$  is the result of the normalization process,  $x_i$  is the data that will be normalized,  $x_{\min}$  is the data with the smallest value,  $x_{\max}$  is the data with the largest value,  $\max$  is the largest value of normalization results which in this case is 1 and  $\min$  is the value of the smallest range of normalized results which in this case is 0. The denormalization process uses the same variables as the normalization process so that each variable has the same function.

## 2.3. Denormalization

After the training process is complete, the normalized values will be returned to their original form to obtain the actual value. This process is called the denormalization process. This process can use Equation (2) [5].

$$x_i = \left( \frac{x' - \min}{\max - \min} \right) * (x_{\max} - x_{\min}) + x_{\min} \quad (2)$$



**Figure 2.** Forecasting system workflow.

#### 2.4. Fitness value

The fitness value is a measure of the performance of a person/flower based on measurements using a certain function [7]. The existence of this fitness value can be a picture of the quality of the flower to be the basis for optimization in the FPA. There is no special function to determine the fitness function. To get an appropriate fitness function, an understanding of the problems being modeled is needed. So each problem being modeled will have a different fitness function.

FPA is an algorithm for optimization. In the optimization case, there are two problems, namely the problem of maximization and the problem of minimization. Both the Genetic Algorithm and FPA work with the concept of maximizing a function  $h$  (maximization), so for the problem of maximizing the fitness value used is the  $h$  function value [7] so that the fitness value follows Equation (3).

$$fitness = \frac{1}{(h + a)} \quad (3)$$

#### 2.5. Flower Representation Design

The implementation of FPA combined with regression will use real-number encoding where each constant will be used as a gene in the flower. The form of representation of interest refers to the form of multiple linear regression equations used in the prediction case shown in Equation (4).

$$Y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad (4)$$

Based on Equation (4) the representation of interest in the FPA will contain the constants used in the regression. Each constant will be represented using 1 gene. so that before the implementation of equation (5) will be changed as follows.

$$Y = \theta_1 + \theta_2 x_1 + \theta_3 x_2 + \dots + \theta_n x_n \quad (5)$$

### 2.6. Sliding Windows Design

Time series data forecasting using FPA will also use sliding windows techniques. This method allows the prediction of periodic data to be more accurate, and the process of predicting currency exchange rates can be determined over time. Sliding windows aim to get new features from the predicted data. Each data will have a different number of sliding windows. The design is as shown in Figure 3.

|                                             |                                             |                                             |                                             |               |               |               |               |               |
|---------------------------------------------|---------------------------------------------|---------------------------------------------|---------------------------------------------|---------------|---------------|---------------|---------------|---------------|
| <b>0.9083</b>                               | <b>0.7963</b>                               | <b>0.6119</b>                               | <b>0.1854</b>                               | <b>0.8919</b> | <b>0.6778</b> | <b>0.7416</b> | <b>0.1483</b> | <b>0.4255</b> |
| Sliding windows 1 ( $\theta_1 - \theta_5$ ) |                                             |                                             |                                             |               | Y             |               |               |               |
|                                             | Sliding windows 2 ( $\theta_1 - \theta_5$ ) |                                             |                                             |               |               | Y             |               |               |
|                                             |                                             | Sliding windows 3 ( $\theta_1 - \theta_5$ ) |                                             |               |               |               | Y             |               |
|                                             |                                             |                                             | Sliding windows 4 ( $\theta_1 - \theta_5$ ) |               |               |               |               | Y             |

**Figure 3.** Sliding windows design.

## 3. Result and Discussion

### 3.1. Analysis of windows size parameters

The windows size parameter influences the accuracy of the results. Windows size provides several inputs which are additional features in forecasting. The process will last for 100 000 iterations. Windows size that will be tested is 3, 5, 6 windows. At the initial determination, 3, 5, 6 are determined and then they can be developed experimentally. There is no specific way of determining window size. The results will be presented in Table 1.

**Table 1.** Windows size training results.

| Windows size | Training      |               | Validation    |               |
|--------------|---------------|---------------|---------------|---------------|
|              | MSE           | RMSE          | MSE           | RMSE          |
| 3            | 0.0044        | 0.0630        | 0.1490        | 0.3720        |
| 5            | <b>0.0043</b> | <b>0.0623</b> | <b>0.0623</b> | <b>0.2186</b> |
| 6            | 0.0068        | 0.0747        | 0.1469        | 0.2919        |

Based on 3 window size experiments, it was found that window size 5 had the smallest MSE. This MSE was deemed sufficient to proceed to the selection of other parameters. In the next parameter test, 5 window sizes will be used. The window size can be adjusted if the MSE is not sufficient or the last prediction results in a bad MSE.

### 3.2. Population size parameter analysis

The population size parameter affects the speed and accuracy of finding solutions. Population size is well correlated with search space, but this is difficult to achieve because of the search space. Get the population size according to the process of experimenting with several population sizes. The population size values to be tested are 15, 20, 30, with the architecture from the previous experiment. There is no definite way of determining population size. The population size is determined

experimentally. Population size initialization uses 15, 20, 30. The total population size can be adjusted according to the needs and results of the MSE. The architecture uses the results of previous experiments where in this case 5 window sizes are used. The test results are in Table 2.

**Table 2.** Population size parameter results.

| Population size | Training      |               | Validation    |               |
|-----------------|---------------|---------------|---------------|---------------|
|                 | MSE           | RMSE          | MSE           | RMSE          |
| 15              | 0.0033        | 0.0233        | 0.2748        | 0.4767        |
| <b>20</b>       | <b>0.0031</b> | <b>0.0116</b> | <b>0.0230</b> | <b>0.1724</b> |
| 30              | 0.0033        | 0.0116        | 0.2164        | 0.402\33      |

Population size 20 produces a better MSE than population size 15 and 30. The MSE value at population size 20 is good enough so that population size 20 will be used for the forecasting process. The population size can be adjusted if the MSE in the forecasting process is not sufficient.

### 3.3. Switch probability parameter analysis

The probability switch affects the ability of the FPA to produce a solution. The probability switch determines when the FPA will intensify or diversify. The probability switches values to be tested are 0.6, 0.7, and 0.8. The probability switch is determined according to the need for a measure of diversification or intensification. Intensification and diversification in FPA are carried out by global or local pollination. So the probability switch size is determined experimentally. The window size used is 5 and the total population is 20. The results obtained are in Table 3.

**Table 3.** Probability switch parameter results.

| Probability switch | Training      |               | Validation |        |
|--------------------|---------------|---------------|------------|--------|
|                    | MSE           | RMSE          | MSE        | RMSE   |
| 0.6                | 0.0011        | 0.0313        | 0.2103     | 0.3634 |
| 0.7                | <b>0.0011</b> | <b>0.0312</b> | 0.0331     | 0.1756 |
| 0.8                | 0.0011        | 0.0331        | 0.2432     | 0.3537 |

Based on the experimental results, the probability switch 0.7 produces the smallest MSE. This MSE is good enough that it will be tested in the forecasting process. So that the forecasting process will use 5 window size parameters, 20 population sizes, and 0.7 probability switch.

## 4. Conclusion

Based on the results of the trials that have been carried out, several things can be concluded as follows.

1. Flower Pollination Algorithm as an alternative to optimizing linear regression can provide a fairly good error value with predictive results close to the actual value. Although this research is still incomplete, the regression forecasting with FPA optimization has shown good results.
2. The architecture of the incorporation of FPA as a linear regression optimization method can be an alternative for determining parameters in linear regression. The search for linear regression parameters can also be performed using other metaheuristic algorithms by following this design.

## References

- [1] Yang X S 2012 Flower pollination algorithm for global optimization *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **7445 LNCS** 240–9
- [2] Yang X S, Karamanoglu M, and He X 2013 Multi-objective flower algorithm for optimization *Procedia Comput. Sci.* **18** 861–8
- [3] Prima F and Trio J 2019 Optimal Reactive Power Dispatch untuk Meminimalkan Rugi Daya Menggunakan Flower Pollination Algorithm *J. Tek. Elektro* **11** 36–43

- [4] Septiawan F Y and Afiahayati 2016 *Prediksi Nilai Tukar Mata Uang dalam Sistem Forex Trading Menggunakan Algoritma Genetika* (Universitas Gadjah Mada)
- [5] Chamidah N, Wiharto, and Salamah U 2012 Pengaruh Normalisasi Data pada Jaringan Syaraf Tiruan Backpropagasi Gradient Descent Adaptive Gain ( BPGDAG ) untuk Klasifikasi **1** 28–33
- [6] Ogasawara E et al 2010 Adaptive Normalization: A Novel Data Normalization Approach for Non-Stationary Time Series *Proc. Int. Jt. Conf. Neural Networks*
- [7] Suyanto 2005 *Algoritma Genetika dalam Matlab* (Yogyakarta: Andi)