**PAPER • OPEN ACCESS**

# Model identification of dengue fever spreading using firefly algorithm and backpropagation neural network

To cite this article: S A Fitania *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **546** 032008

View the article online for updates and enhancements.

# **IOP ebooks**™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

# Model identification of dengue fever spreading using firefly algorithm and backpropagation neural network

**S A Fitania[1], A Damayanti[2,*], and A B Pratiwi[3]**

[1,2,3]Department of Mathematics, Faculty of Science and Technology, Airlangga University, Indonesia

*Corresponding author: aulid@fst.unair.ac.id[2]

**Abstract.** Dengue Fever is one of Indonesia's well-known medical problems where the range spread territories have became more extensive alongside with mobility and population growth. Considering that a large number of population in East Java - Indonesia has been infected, the identification of Dengue Fever is needed in order to anticipate and minimalize the terrible possibilities that could happen. The aim of this research is to obtain the result of Dengue Fever spreading model identification using Firefly Algorithm and Back-propagation Neural Network. Back-propagation Neural Network identification is proposed to estimate the spreading of Dengue Fever based on actual data. The process begins with estimating the parameters using Firefly Algorithm then identifying the model using Back-propagation Neural Network. Based on the implementation and simulation on the Dengue Fever spreading data in East Java - Indonesia from January 2013 to December 2017, model was succesfully identified where the error value between estimated data and actual data was 0.0242.

## 1. Introduction

Fever is a condition of the body temperature above normal as a result of temperature control center increase in the hypothalamus [1]. One of the diseases showing fever symptoms is Dengue Fever. Dengue Fever is a diseases caused by *Aedes Aegypti* mosquitos [2]. Dengue Fever is one of the Indonesia's well-known medical problem where the total and range spread territories that are getting bigger and wider alongside with mobility and growth of the citizen population [3]. In Indonesia, Dengue Fever was discovered in East Java province in 1968. There were 58 infected individuals and approximately 41.3% of them could not survive. Since then, this particular disease has spread extensively across Indonesia. The highest percentages of group infected by dengue fever in Indonesia at 5-14 years old and 15-44 years old range of ages were 43.44% and 33.25%, respectively [3]. Considering that Dengue Fever has infected individuals in East Java in large amounts therefore an identification of Dengue Fever is needed in order to anticipate and minimalize terrible possibilities that can happen caused by a massive number of Dengue Fever infected fluctuation.

Identification is organizing certain system hence the output corresponds the target. Identification can be solved using several methods, one of them is Artificial Neural Network (ANN) [4]. ANN is one of the artificial representation of human brain working to simulate the process of human brain works [5]. ANN can be applied intensively on forecasting, especially sales prediction. Many research concludes that ANN method is better in performance than forecasting conventional methods [6]. ANN also has data mining process in order to obtain optimal weights matched with the training pattern. ANN is very effective to identify a complicated system [7].

Many ANN methods have been proposed to identify system, and one of them is Back-propagation (BP). BP Algorithm has an upside part of it where the interconnection weights can give simulation results with good accuracy [8]. To help the process of identification, therefore Firefly Algorithm (FA) is proposed as a method to estimate the parameters before identification process using BP. According [9], FA is an algorithm inspired by the behavior of fireflies in flashing. Each firefly will be attracted and moves toward other brighter firefly. According [10], all variables constructed by firefly movement equations are guaranteed can make the algorithm succeed in finding the optimal solution.

Based on the explanation above, this paper discuss about the implementation of Firefly Algorithm with Back-propagation to identify the spreading of Dengue Fever. This paper is organized as follows: next section discusses the methodology, section 3 provides result and discussion, and the conclusion is presented in last section.

## 2. Methodology
The steps taken in this study include determining model and data, estimating parameters using Firefly Algorithm, model identification using Backpropagation neural network and model validation.

### 2.1 Model and Data
Model identification of the dengue fever spreading refers to a journal written by Pandey et al on 2013. The model contains three compartments which are $S$ compartment (susceptible human population), $I$ compartment (human population infected by dengue fever) and $R$ compartment (recovered human population from dengue fever). Dengue fever spreading data was collected from East Java's Department of Health 2018 from January 2013 to December 2017 which contained total of healthy individuals, total of dengue fever infected individuals and total of dengue fever recovered individuals [11]. A standard SIR model is given as follows:

$$\frac{dS}{dt} = B - \omega \frac{I}{N} S - \mu S \tag{1}$$

$$\frac{dI}{dt} = \omega \frac{I}{N} S - \varepsilon I - \mu I - \theta I \tag{2}$$

$$\frac{dR}{dt} = \varepsilon I - \mu R \tag{3}$$

where:
$B$: number of human birth
$\omega$: transmission rate of disease between human
$\mu$: rate of human death
$\varepsilon$: rate of human recovery
$\theta$: death rate caused by disease

Total of population is $N = S + I + R$. Because the number of population is positive, it is assumed that $S, I, R, B \geq 0$. In addition, the value of rate is a positive scalar quantity. Then all parameters are assumed to be positive on interval (0,1).

### 2.2 Parameters Estimation
Firefly Algorithm is used to estimate parameters of the model. Firefly Algorithm is algorithm that is inspired by the pattern movement of fireflies, where every firefly will have an interest to approach fireflies with brighter light **[9]**. Firefly Algorithm process in order to estimate the parameters of Dengue Fever spreading model are input data and initialize parameters, generate firefly population, calculate numeric solution and calculate the error using Mean Magnitude of Relative Error ($MMRE$), calculate the brightness value and compare the brightness value of each firefly and then move to firefly with brighter light. Flowchart of Firefly Algorithm is presented in Figure 1.

Two important procedures in parameters estimation are listed as follows:

1. Calculate Numeric Solution
   The procedures of numeric solution calculation for each firefly using Runge-Kutta four orders method are listed below:

   i. Determine dengue fever mathematic model that will be solved numerically Runge-Kutta four orders method [12].

   ii. Solve dengue mathematics model numerrically using Runge-Kutta four orders method [12].

2. Calculate MMRE

Mean Magnitude of Relative Error (MMRE) [13] is obtained from calculation result of numeric solution using Runge-Kutta four orders method and actual data.

$$MMRE = \frac{1}{n}\left(\sum_{i=1}^{n}\frac{|y_i - y_i^*|}{y_i}\right) \tag{4}$$

where $n$ is number of data, $y_i$ is actual data value and $y_i^*$ calculated data value.
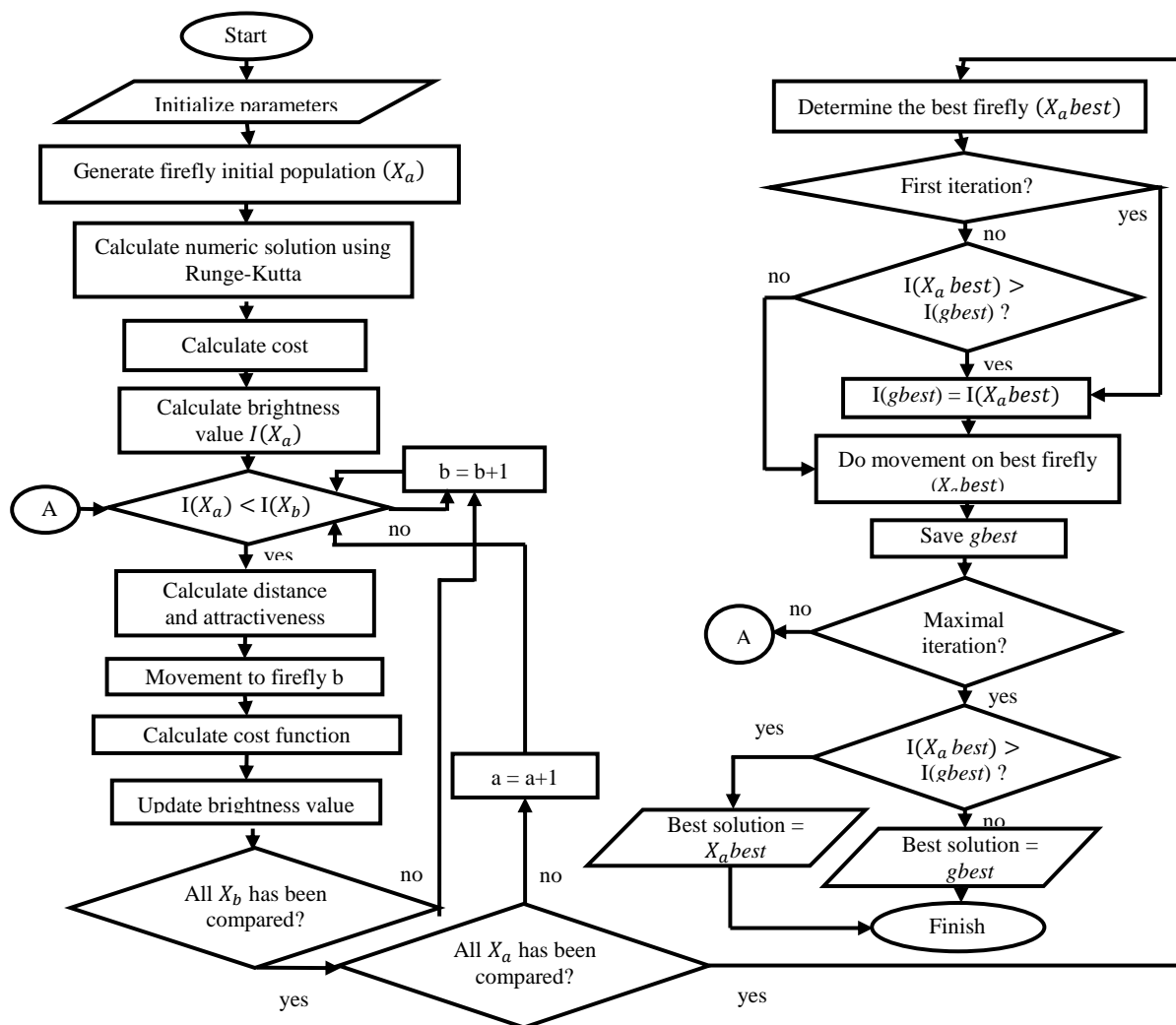


**Figure 1.** Flowchart of FA Process

*2.3 Model Identification*

Procedures of model identification using Backpropagation neural network are initialize parameters, normalization process and backpropagation process. The procedures are explained as follows:

1. Initialize Parameters
   Parameters initialization including the weights and biases which will be used in Backpropagation process.

2. Normalization Process
   Apply normalization process so that the Dengue Fever data is converted to interval (-1,1) using equation below.

$$x = \frac{2\,(x_p - \min(x_p))}{\max(x_p) - \min(x_p)} - 1 \tag{5}$$

   where $x$ is data normalization value, $x_p$ is data value before normalization process, $\min(x_p)$ is minimum data value before normalization process, and $\text{maks}(x_p)$ is maximum data value before normalization process.

3. Back-propagation Process
   Back-propagation consists of feed forward and back-propagation of error process. Back-propagation neural network architecture can be seen in Figure 2 [5].
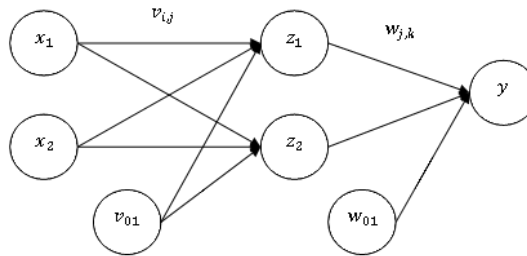


**Figure 2**. Back-propagation Neural Network Architecture

When feedforward occurs, every unit $(x_i)$ will receive input signal and will spread the signal on every hidden unit $(z_j)$. Every hidden unit is going to calculate the activities and send signal $(z_j)$ to every output unit. Then, every output unit $(Y_k)$ also calculate the activities value $(y_k)$ that also give a respond to input that has been given by network.

During the training process, every output unit is compared to the activities $(y_k)$ with targeted value $(t_k)$ the determine the number of error. Based on this error, factor $\delta_k$ $(k = 1, \ldots, m)$ can be calculated. $\delta_k$ is used to distribute error on output unit $Y_k$ back to every unit on previous layer. This is also used to update the weight between output and hidden layer. With the same method, factor, $\delta_j$ $(j = 1, \ldots, p)$ is calculated on every hidden unit $Z_j$. Factor $\delta_j$ is used to update the weight between hidden layer and input layer.

After every factor $\delta$ is determined, weight from every layer is updated together. Weight update $w_{jk}$ is based on $\delta_k$ factor and activation $z_j$ from hidden unit. Meanwhile updated weight $v_{ij}$ is based on $\delta_j$ factor and activated $x_i$ from input unit. The number of input units used in this paper is two, as well as hidden units. Whereas for the output unit used is one. Activation function that used in this paper is sigmoid bipolar:

$$y = f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \tag{6}$$

With differential of the activation function:

$$f'(x) = [1 + f(x)]\,[1 - f(x)] \tag{7}$$

where $y$ is output value and $x$ is input value.

### 2.3.1 *Feedforward*

The steps of the feedforward process are explained as follows:

i.  Each hidden unit is sum of input signal with its weight and bias:

$$z_{in_j} = v_{0j} + \sum_{i=1}^{n} x_i v_{ij} \tag{8}$$

And then using the activation function that is used to calculate the output signal and hidden unit:

$$z_j = f\left(z_{in_j}\right) \tag{9}$$

Then send the output signal to the entire output unit.

ii. Each output unit is sum of the input signal with its weight and bias:

$$y_{in_k} = w_{ok} + \sum_{j=1}^{p} z_j w_{jk} \tag{10}$$

And using the activation function to calculate output signal and output unit:

$$y_k = f\left(y_{in_k}\right) \tag{11}$$

### 2.3.2 *Back-propagation of Error*

The steps of the back-propagation of error process are as follows:

i.  Each output unit receives certain target that is qualified as a fit to training input data to calculate error between target and output:

$$\delta_k = (t_k - y_k)f'\left(y_{in_k}\right) \tag{12}$$

As the input data of training, output data of training $t_k$ has also been scaled according to activation function that is currently used. Factor $\delta_k$ is used to calculate the error correction $\left(\Delta w_{jk}\right)$ that is used to update the weights, $w_{jk}$:

$$\Delta w_{jk} = \alpha \delta_k z_j \tag{13}$$

And also, $\Delta w_{0k}$ , weight correction is calculated to be used to update weights, $w_{ok}$:

$$\Delta w_{ok} = \alpha \delta_k \tag{14}$$

ii. Each hidden unit sums the delta input*:*

$$\delta_{in_j} = \sum_{k=1}^{m} \delta_k w_{jk} \tag{15}$$

Then multiplied with differential of the activation function to calculate the error information:

$$\delta_j = \delta_{in_j} f'\left(z_{in_j}\right) \tag{16}$$

Then calculate weights correction:

$$\Delta v_{ij} = \alpha \delta_j x_i \tag{17}$$

The bias weights correction, $\Delta v_{0j}$, is also calculated to update $v_{oj}$.
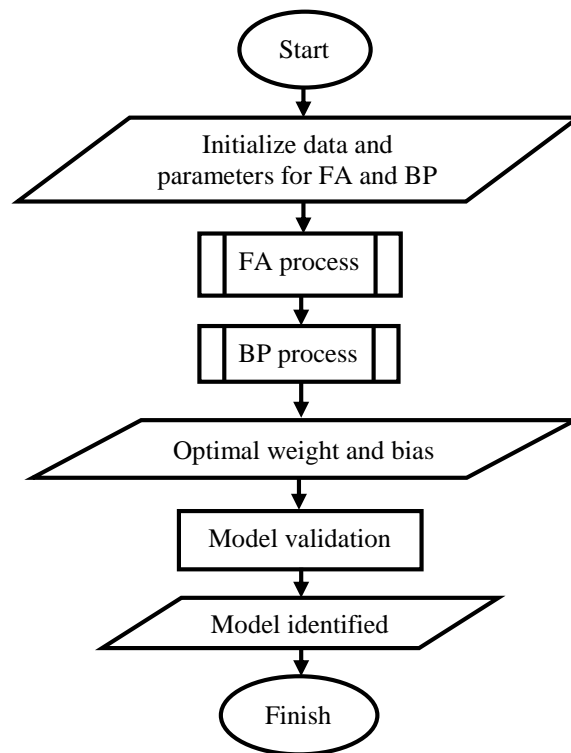
$$\Delta v_{oj} = \alpha \delta_j \tag{18}$$

iii. Each output unit updates the weights and bias:
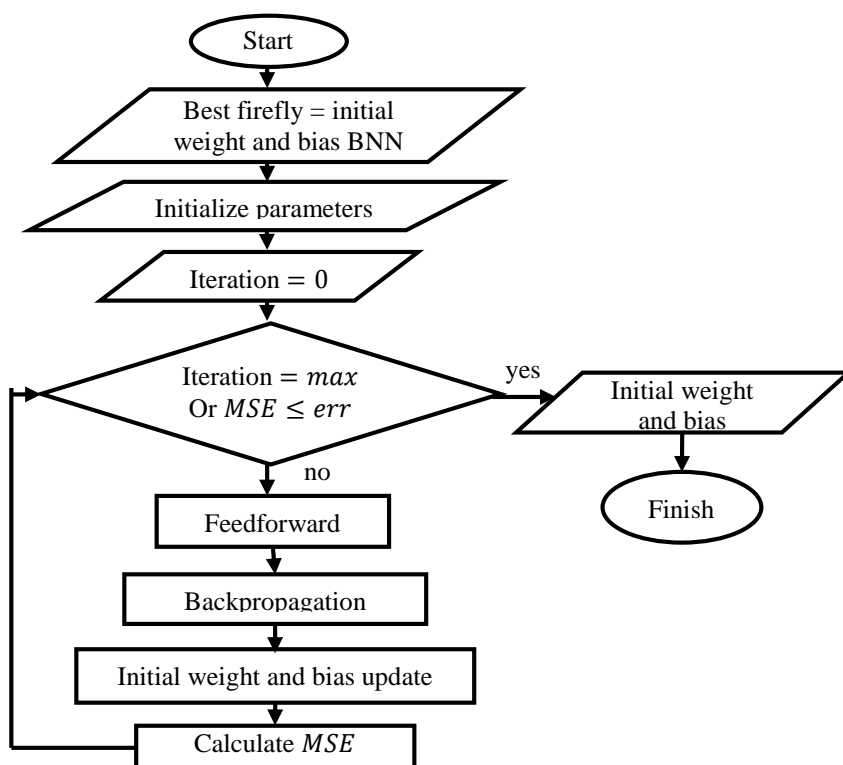
$$w_{jk}(new) = w_{jk}(old) + \Delta w_{jk} \tag{19}$$

And also, each input unit updates the weights and bias:

$$v_{ij}(new) = v_{ij}(old) + \Delta v_{ij} \tag{20}$$

The steps taken in the identification model method are shown in Figure 3 and the detail steps of Backpropagation neural network process are shown in Figure 4.

**Figure 3.** Flowchart of Identification Model



**Figure 4.** Flowchart of Backpropagation Neural Network Process

*2.4 Model Validation*

The next step is applying the validation model process which is feedforward process using the optimal weights and bias gained from previous process. Then calculate error differences between back-propagation neural network calculation and actual data.

## 3.  Results and Discussions

The data used in this study are 180 data collected from East Java's Department of Health (2018) from January 2013 to December 2017 which is consisted of 60 data of healthy individuals, 60 data of dengue fever infected individuals and 60 data of dengue fever recovered individuals [14]. Sample of data used in this study can be seen in Table 1.

**Table 1.** Data of Dengue Fever Spreading in East Java Province 2013-2017

| No. | Month | Total of Healthy Individuals | Total of Dengue Fever Infected Individuals | Total of Dengue Fever Recovered Individuals |
|---|---|---|---|---|
| 1 | January 2013 | 38698320 | 3272 | 3246 |
| 2 | February 2013 | 38699775 | 2549 | 2514 |
| 3 | March 2013 | 38700723 | 2067 | 2048 |
| 4 | April 2013 | 38701244 | 1803 | 1791 |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| 59 | November 2017 | 39506389 | 494 | 487 |
| 60 | December 2017 | 39506255 | 562 | 553 |

In this study, there are three steps that must be done, including parameter estimation, model identification and model validation. In the first process, which is estimated parameter, after obtaining the optimal model parameters from this process, the parameters will be converted into initial weights in the next process, namely model identification. The model identification process aims to produce optimal weights and biases that will be used for the next process, which is model validation. In the process of model identification using 60 data in each population. The accuracy of model identification using Firefly Algorithm and Backpropagation neural network for $S$, $I$ and $R$, can be seen on Figure 5, Figure 6 and Figure 7, respectively. Figure 5 shows that the simulation result of model identification for suspectible population gives good accuracy in identifying the actual data, while the pattern trends of infected population can be identified well at the some period of time. Similarly, the recovery population is presented in Figure 7.  Table 2 explains the validation result using some variation in learning rate of backpropagation. After 1000 iteration, the best validation error (0.0242) is obtained when the learning rate is 0.723. It is also shown that the value of Mean Square Error (MSE) for the process is 3.6 x $10^{-5}$.
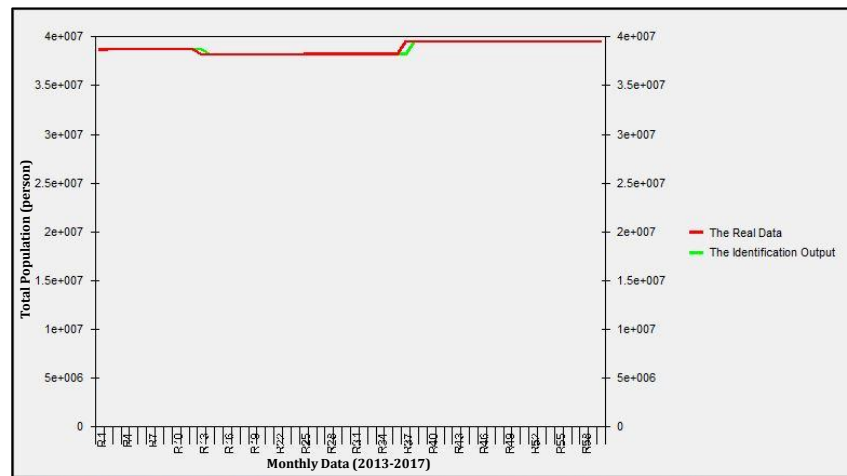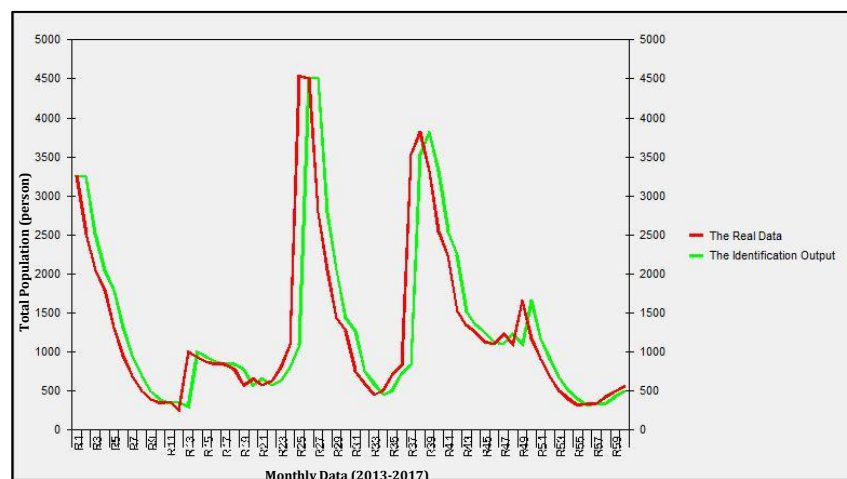
**Figure 5.** Result of Model Identification for *S* population
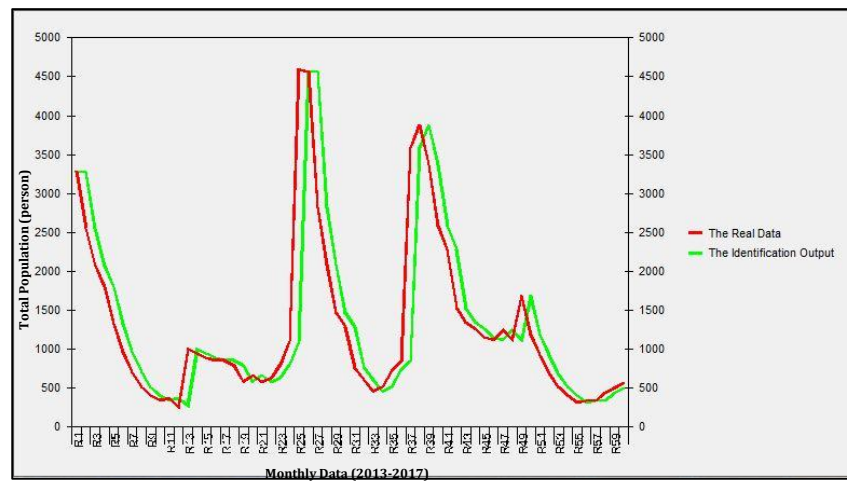


**Figure 6.** Result of Model Identification for *I* population



**Figure 7.** Result of Model Identification for *R* population

**Table 2.** Identification using Various Learning Rates

| Experiment | Learning rate | MSE | Validation Error |
|---|---|---|---|
| 1 | 0.1203 | $3.7 \times 10^{-5}$ | 0.0285 |
| 2 | 0.7007 | $3.7 \times 10^{-5}$ | 0.0275 |
| 3 | 0.723 | $3.6 \times 10^{-5}$ | 0.0242 |
| 4 | 0.072 | $4.1 \times 10^{-5}$ | 0.0248 |
| 5 | 0.256 | $3.9 \times 10^{-5}$ | 0.0258 |
| 6 | 0.0683 | $3.8 \times 10^{-5}$ | 0.0278 |
| 7 | 0.9252 | $4.0 \times 10^{-5}$ | 0.0268 |
| 8 | 0.3037 | $3.9 \times 10^{-5}$ | 0.0245 |
| 9 | 0.6431 | $3.8 \times 10^{-5}$ | 0.0269 |
| 10 | 0.8311 | $3.8 \times 10^{-5}$ | 0.0268 |

## 4. Conclusion

The purpose of this paper is to identify of Dengue Fever spreading model using Firefly Algorithm and Back-propagation Neural Network. The procedures of identification Dengue Fever spreading model are parameters estimation using Firefly Algorithm, then identification and validation model using Backpropagation neural network. Based on the simulation results, shows that the system is able to identify the data patterns with an error rate 0.0242.

## References

[1]   Sodikin, 2012, *Prinsip Perawatan Demam Pada Anak,* Pustaka Belajar, Yogyakarta.

[2]   Gubler D.J., 1998, *Dengue and dengue hemorragic fever.* Clinical Microbiology Reviews, **11**(3): 480-496.

[3]   Kementerian Kesehatan RI, 2010, *Demam Berdarah Dengue di Indonesia Tahun 1968-2009,* Buletin Jendela Epidemologi, **2**:1.

[4]   Rubio, J.J., Yu, dan Wen, 2007, *Nonlinier system identification with recurrent neural network and dead-zone Kalman filter algorithm,* Science Direct, Neurocomputing, **70**:2460-2466.

[5]   Fausett, L., 2003, *Fundamental of Neural Network Architecture, Algorithm, and Application,* Printice-Hall Inc. London.

[6]   Sun, Z.L., Choi, T.M., Au, K.F., dan Yu, Y., 2008, *Sales Forecasting using Extreme Learning Machine with Application in Fashion Retailing*, Decision Support Systems, **46**: 411-419.

[7]   Haykin, S., 1999, *Neural Network A Comprehensive Foundation Second Edition*, in New Jersey Prentice-Hall Inc.

[8]   Brajevic, I. dan Tuba, M., 2013, *Training feed-forward neural networks using firefly algorithm,* Recent Advances in Knowledge Engineering and Systems Science, pp. 156-161, WSEAS.

[9]   Yang, X, S., 2010, *Firefly Alforithm, Levy Flights And Global Optimization.* In Research and Development in Intelligent Systems XXVI, pp. 209-218, Springer, London, UK.

[10]  Broersma, H., 2010, *Application of the Firefly Algorithm for Solving the Economic Emmisions Load Dispatch Problem*, International Journal of Combinatorics, Hindawi Publishing Corporation.

[11]  Pandey, A., Mubayi, A., Medlock, J., 2013. *Comparing vector-host and SIR models for dengue transmission,* Mathematical Biosciences, **246**:252-259.

[12]  Zill, D.G., Cullen, M.R., 2009, *Differential Equations With Boundary-Value Problem, seventh edition,* Nelson Education Ltd: Canada.

[13]  Foss, T., 2003, *A simulation study of the model evaluation criterion MMRE*, *IEEE Transactions on Software Engineering*, **29**(11): 985-995.

[14]  Dinas Kesehatan Provinsi Jawa Timur, 2018, *Profil Kesehatan Provinsi Jawa Timur tahun 2013-2017,* Dinas Kesehatan Provinsi Jawa Timur, Surabaya.