Mustaffa et al. *Iraqi Journal of Science, 2020, Vol. 61, No. 8, pp: 2132-2141*
DOI: 10.24996/ijs.2020.61.8.28

ISSN: 0067-2904

An application of Barnacle Mating Optimizer in Infectious Disease Prediction: A Dengue Outbreak Cases

Zuriani Mustaffa[1*], Mohd Herwan Sulaiman[2], Mohamad Farhan Mohamad Mohsin[3], Yuhanis Yusof[3], Ferda Ernawan[1], Bariah Yusob[1], Noorhuzaimi Mohd Noor[1]
[1]Faculty of Computer Systems & Software Engineering, Universiti Malaysia Pahang, 26300 Kuantan, Pahang, Malaysia
[2]Faculty of Electrical & Electronics Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia
[3]School of Computing, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia

Received: 21/8/2019 Accepted: 19/11/2019

Abstract

Meta-heuristic algorithms have been significantly applied in addressing various real-world prediction problem, including in disease prediction. Having a reliable disease prediction model benefits many parties in providing proper preparation for prevention purposes. Hence, the number of cases can be reduced. In this study, a relatively new meta-heuristic algorithm namely Barnacle Mating Optimizer (BMO) is proposed for short term dengue outbreak prediction. The BMO prediction model is realized over real dengue cases data recorded in weekly frequency from Malaysia. In addition, meteorological data sets were also been employed as input. For evaluation purposes, error analysis relative to Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE), and Mean Absolute Deviation (MAD) were employed to validate the performance of the identified algorithms which includes the comparison between BMO against Moth Flame Optimizer (MFO) and Grey Wolf Optimizer (GWO) algorithms. Upon simulation, the superiority is in favour to BMO by producing lower error rates.

Keywords: Barnacles Mating Optimization, Dengue outbreak prediction, Meta-heuristic, Optimization, prediction

Introduction

Due to the outstanding performance of meta-heuristic algorithms in addressing various practical problems [1], a good number of researches have been published which provides powerful contributions to the interested party. The glorious of meta-heuristic era begins with the introduction of well-regarded Genetic Algorithm (GA) [2] which simulates the natural selection and genetic mechanism of Darwin's biological theory. Given the growing number of GA applications, a good number of publications can be seen across different fields which include face recognition [1], electrical [3], cloud computing [4], operation management [5] and many more. This proves that GA has contributed a significant impact in addressing many real-world problems. Progressing further, meta-heuristic algorithms were later become surprisingly popular with the present of many animals inspired based algorithms such as Particle Swarm Optimization (PSO) [6] which mimics the collective behaviour of birds in navigating and Ant Colony Optimization (ACO) [7] which simulates the social behaviour of ants in finding for the shortest food source. Unlike GA which is classified as Evolutionary Algorithm (EA), PSO and ACO are categorized as Swarm Intelligence (SI) algorithm. Other than PSO and ACO, recently the family of SI algorithms become bigger with Grey Wolf Optimization (GWO) [8], Moth-flame Optimization (MFO) [9], Salp Swarm Algorithm (SSA) [10]

---

*Email: zuriani@ump.edu.my

and Barnacles Mating Optimizer (BMO) [11], which is one of the most newly introduced SI algorithm. The application of SI algorithms are well accepted in numerous area and this includes in infectious disease prediction [12] [13].

Among of the critical infectious disease, dengue fever which is an infectious disease is continuously recognised as a serious, worldwide public health concern. Presented in more than 126 countries, it is known as the most prevalent mosquito borne disease in the world, which leading to 390 million of dengue infections annually, all over the globe [14]. The only continent that does not affected with dengue transmission are Europe and Antarctica [15]. Dengue fever carries similar symptoms to other fever such as Viral influenza, Chikungunya, Zika fever, and a few others [12]. Once positively diagnosed, the virus could cause serious reduction of blood platelets, which could lead to fatal.

In this case, Malaysia is not excluded. It is reported that the most affected regions are central region which include Selangor, Kuala Lumpur and southern region as well, which is Johor. With the increasing number of cases each year, it is rapidly becoming as the most critical health concern. It is recorded that a huge increment of number of dengue cases was reported between 2000 to 2014 with 91% of increments, from 32 cases per 100,000 population in year 2000 to 361 cases per 100,000 population in 2014 [16]. With such alarming situation, tremendous effort have been done to find the most effective vaccine to combat dengue, which includes the vaccine development, but to no avail [15].

The remaining sections are structured as follows: The next section provides the existing works on the issue of interest, followed by description on BMO as a predictor model. Next, experimental setup of this work is discussed which covers data collection, data scaling, evaluation and parameter setting. Meanwhile, the following section presents the obtained results and finally the main conclusions of this work is drawn on the last section.

Related Work

Concerning this matter, quite number of studies related to prediction and early detection of dengue have been published, which witnessed various techniques have been presented, with no exception on the application of meta-heuristic based algorithms. A hybrid model of Particle Swarm Optimization–Artificial Neural Network (PSO-ANN) has been proposed for earlier diagnosis of dengue disease [12]. In the proposed model, the PSO served as an optimizer to optimize the weight and bias parameters of ANN, before the detection task takes place. For experiment purposes, the proposed method was tested on real data of dengue from Delhi region. Based on a few metrics recorded which include the accuracy, the PSO-ANN is proven to be proficient and powerful model for the case under study.

Previously, an application of ANN for dengue confirmed cases has been presented by using physical parameters of mean temperature, mean relative humidity and total rainfall [17]. The results indicated that stated features including the total number of dengue confirmed-cases were efficient in predicting the number of dengue confirmed-cases. Recently, findings for dengue outbreak prediction using four hybrid meta-heuristic algorithms have been demonstrated, which limited to the data collected from Yogyakarta, Indonesia [13]. The obtained results indicate that the presented hybrid models are able to produce competitive results, with a slightly favour to Artificial Bee Colony-Least Squares Support Vector Machines (ABC-LSSVM) model. Similar studies also can be seen in [18].

Earlier in 2012, a hybrid dengue outbreak prediction model based on GA-ANN has been proposed [19]. The GA-ANN has been realized on time series of dengue cases data, rainfall data and proximity location data which were collected from State Health Department of Selangor (SHD) and Malaysian Meteorological Department. In the study, GA is employed for the determination of weight in neural network model. The findings show that the proposed hybrid model is superior compared to standalone models.

Despite numerous published studies on this issue, the importance of proper dengue outbreak cases prediction model is still wide open for an improvement as this issue involved human's life. Concerning this matter, this study aims to propose a dengue outbreak cases prediction model based on BMO algorithm. Despite the young age, BMO has been proven to outperform other meta-heuristic algorithm which includes MFO[11]. Later, the effectiveness of the BMO prediction model will be validated against comparable meta-heuristic algorithm namely GWO and MFO.

Methodology of BMO as a predictor

This section provides an insight on BMO. The first subsection will give a brief description on barnacles in nature, followed by the mathematical model of BMO.

Barnacles Mating Optimizer in nature

Deep in the ocean, barnacles survive with such a unique social life. They commonly found permanently attached to rocks, corals or ships. Governed by its one of a kind features where its penises are probably the longest among microorganism, relative to its body, they perform mating activity to overcome their sessile state.

A. Mathematical Model

i) Initialization

The candidate solutions are represented as in (1):

$$X = \begin{bmatrix} x_1^1 & \cdots & x_1^N \\ \vdots & & \vdots \\ x_n^1 & \cdots & x_n^N \end{bmatrix} \tag{1}$$

Where:

N = number of control variables to be optimized

n = number of population/barnacles

Similar with other existing meta-heuristic algorithms, the BMO begins with initialization where the candidate solutions are randomly generated. In order to obtain the initial performance of the candidates, an evaluation will be done, before the sorting process takes place.

Selection

Selection process is one of the critical stages in BMO. For this matter, the following rules are provided:

i. The selection is a random process which is subject to the constraint *i.e* length of the penis, *pl.*

ii. In nature, every barnacle contributes and receives its sperm from other barnacles due to hermaphroditic behavior. However, in BMO, at one particular time, it is assumed that each barnacle can fertilized only by one barnacle even though in [20], it is noted that more than one male can copulate to a single female barnacle. This is implemented for the sake of simplicity.

iii. At the certain iteration, if the algorithm selects the same barnacle, the self-mating process assumed to be happened. Nevertheless, self-mating is very rare [21]. Thus, it will not be considered in the proposed BMO. When the stated situation occurred, no new offspring will be produced.

iv. If at certain iterations, the selection for mating is out of range of *pl* that has been set, it is assumed that the new off-springs are produced by sperm-cast mating process

ii) Off-spring generation

The generation of new off-spring obeys the Hardy-Weinberg equilibrium principle [22]. Like any other meta-heuristic algorithms, two important processes namely exploitation and exploration do take place n BMO algorithm. The exploitation process is defined as follow:

$$x_{child}^{N\_new} = p x_{dad}^{N} + q x_{mum}^{N} \; for$$

$$x_{mum}^{N} \qquad k \leq pl \tag{2}$$

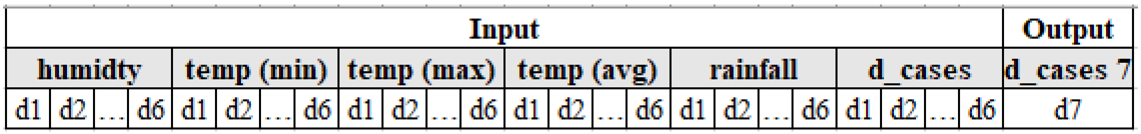
Meanwhile, (3) shows the definition of exploration process:

$$x_{child}^{N\_new} = [rand() \times x_{mum}^{N}] + x_{mum}^{N} \; for \; k > pl \tag{3}$$

Where p represents the normally distributed random numbers, q=(1-p), x_{dad}^{N} and x_{mum}^{N} indicate the variables of *Dad* and *Mum* of barnacles respectively, which has been chosen for mating. Meanwhile, *rand*() is the randomly generated number between the range of 0 to 1 and *k* denotes the range of barnacle’s penis can be reached during mating.

B. BMO prediction model

In this study, BMO algorithm will be served as an estimator of parameters’ vector. The inputs consist of humidity, minimum, maximum and average of temperature, rainfall and the number of dengue cases. The derivation of the stated inputs until day 6 are fed to the prediction model as well, as illustrated in Figure-1. Meanwhile, the output is the dengue cases from day 7 and onwards.



Input																		Output
humidty			temp (min)			temp (max)			temp (avg)			rainfall			d_cases			d_cases 7
d1	d2…	d6	d1	d2…	d6	d1	d2…	d6	d1	d2…	d6	d1	d2…	d6	d1	d2…	d6	d7

** d =day, temp = temperature, min = minimum, max = maximum, avg = average, d_cases = dengue cases, d_cases 7 = number of dengue cases from day 7 and onwards*

Figure 1- Input Output Configuration

Experimental Setup

A. Description of experimental data

This study employed a real time series data of dengue cases collected from several districts in Peninsular Malaysia. The data were recorded in weekly frequency, from 2003 to 2008 incorporated of weekly dengue cases incidence data as well as meteorological data namely humidity, minimum, maximum and average value of temperature and rainfall values. Table-1 shows the sample of the employed data.

Table 1- Sample of Employed Data

Year	**Weeks**	**Humidity (g/m3)**	**Temperature °c**			**Rainfall value (mm)**	**Dengue cases incidence**
			Min	**Max**	**Avg**		
2005	10	66.29	24.67	34.26	28.34	24.14	19
	11	65.33	24.79	36.01	29.05	32.41	15
	12	72.18	24.03	34.91	28.22	55.66	9
	13	77.03	23.57	32.96	27.26	40.2	13
	14	75.99	23.24	33.36	27.27	130.13	19

**Min=minimum, max=maximum, avg=average*

The data were later split into training and testing sets, each consist of 70% and 30% of the observations, respectively.

B. Data Scaling

Prior to training, the data were first scaled to deal with out-of-bound attributes [23]. This is important to reduce the complexity of the employed data. In this study, the data were scaled using Min-Max normalization, which is defined as follow:

$$v' = \frac{v - \min_A}{\max_A - \min_A} \tag{5}$$

Where v and v' indicate original value and normalized value respectively, while $\min_a$ and $\max_a$ represent the minimum and maximum value of the feature A.

C. Performance metrics

Three quantitative indices were employed for evaluation purposes, namely Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE), and Mean Absolute Deviation (MAD). The definitions of the stated metrics are shown in (6), (7) and (8).

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - Y_i|}{y_i}.100 \tag{6}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - Y_i)^2 \tag{7}$$

$$MAD = \frac{1}{n}\sum_{i=1}^{n}|y_i - Y_i| \tag{8}$$

Where y_i denotes the actual value for the *i*th data point, Y_i indicates the predicted value at period *i*, and n is the number of prediction period.

D. BMO parameter setting

Before the experiments take place, the properties of identified are defined as follow: number of population =30, maximum iteration=100, and *pl* (BMO parameter) = 25.

Empirical results and discussion

The experiment is executed using MATLAB platform on Intel®Core™ i3-5005U processor, CPU @ 2.00 GHz with 4.00 GB of RAM, in 64-bit Windows 10 operating system. The experiment is divided into two stages, where the first stage was conducted to compare the performance of BMO prediction model using raw and normalized data (see section 4.2). The BMO were trained using approximately 3 years and half of weekly recorded data, from 2004 – 2007, and later were tested on 2007-2008 of data, producing predictions of dengue outbreak cases for one week ahead. The obtained results are shown in Table- 1.

Table 1-Prediction Evaluation for BMO prediction model: Raw vs. Normalized data

	Raw data	**Normalized data**
MAPE (%)	27.7082	27.3892
MSE	94.7667	90.8693
MAD	7.0506	6.9609

Based on the table, it is demonstrated that the BMO was able to produce lower error rates when realized on normalized data where the MAPE produced is 0.319% lower than MAPE produced when the model is fed with raw data. The other two metrics namely MSE and MAD showed a good agreement as well.

Data tabulated in Table-2 showed prediction values generated by BMO by using the stated data during testing period, from week 265 until 279. The bold font indicates the closest value produced by BMO, on respective data set. As shown in the table, the BMO is capable to predict closer to target values when realized on normalized data, especially in week 273.

Table 2- Prediction value for BMO from week 265 to 279: Raw vs. Normalized data

Week	**Target**	**Raw Data**	**Normalized Data**
265	19	22.2398	**21.5417**
266	18	**18.3264**	16.8084
267	21	13.594	**16.0785**
268	30	14.6165	**15.4803**
269	15	**19.4453**	19.6312
270	18	**15.701**	14.0968
271	26	12.4055	**15.3338**
272	38	16.7796	**17.1834**
273	23	24.4164	**23.6493**
274	16	21.5889	**19.0316**
275	31	13.959	**16.1934**

276	30	18.3378	**20.3135**
277	30	**22.4836**	20.2301
278	21	22.2119	**22.1333**
279	23	**18.2732**	17.9749

The comparison performance of BMO prediction model when tested on raw and normalized data is visualized in Figure-2. By referring to the figure, blue solid line indicates the target value, green dashed-dotted line represents the raw data while red dashed line denotes prediction values produced by BMO using normalized data.

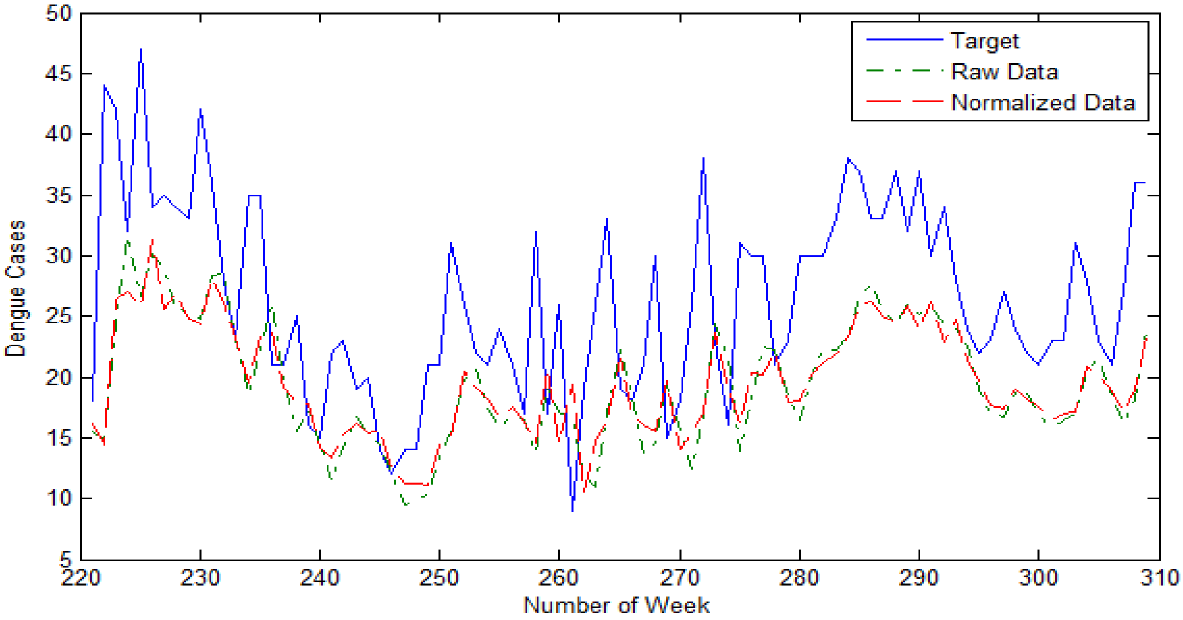


Figure 2- BMO Prediction Model for Dengue Outbreak: Target vs. Raw Data vs. Normalized Data

Meanwhile, the convergence value produced by BMO by using raw and normalized data is illustrated in Figure-3. Based on the figure, the BMO converged at a very similar values, which are 0.0253 when using raw data and 0.0256 when tested on normalized data.

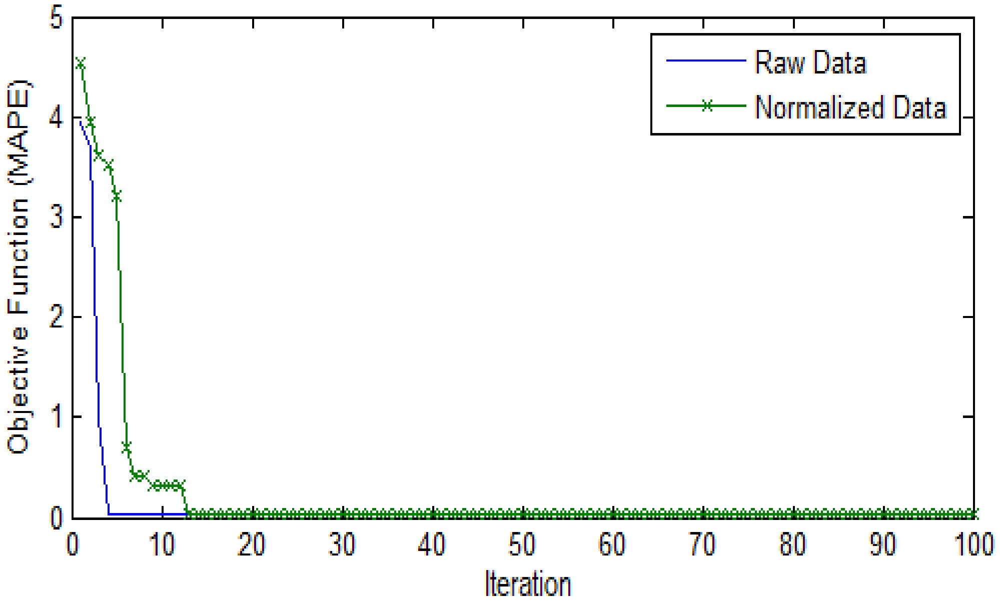


Figure 3- Convergence rate for BMO: Raw data vs. Normalized data

The second stage of the experiment is to proceed with normalized data by performing a comparison between BMO and other identified prediction model, namely MFO and GWO. The findings of the experiment are tabulated in Table-3. By referring to the table, the prediction values of each of the 3 identified models are tested against each other over an approximately 2-years period. The BMO prediction model is able to produce lowest error rates, where the MAPE, MSE and MAD are 27.3892%, 90.8693 and 6.9609 respectively. Meanwhile, the MFO which is ranked second is able to record 24.6111% lower MAPE compared to GWO, which 33.66%. Meanwhile, the MSE and MAD recorded by MFO are 122.8911 and 8.8820 respectively.

Table 3-Prediction Evaluation for BMO vs. MFO vs. GWO using normalized data

	BMO	**MFO**	**GWO**
MAPE(%)	27.3892	33.66	58.2711
MSE	90.8693	122.8911	310.2388
MAD	6.9609	8.8820	16.1967

Table-4 shows the prediction values recorded by all three identified prediction models, from week 265-279. For the convenience of comparison, the closest values produced by the respective prediction model are bolded. The visualized results are also provided, as shown in Figure-4. By referring to the figure, solid blue line represents the actual value while the results produced by BMO is indicated by dashed green line. Meanwhile, crossed and dotted marks denote prediction values recorded by MFO and GWO respectively.

Table 4- Prediction value for dengue outbreak cases using normalized data from week 265 to 279: BMO vs. MFO vs. GWO

Week	**Target**	**BMO**	**MFO**	**GWO**
265	19	21.5417	**21.37**	11.71
266	18	**16.8084**	13.40	8.82
267	21	**16.0785**	12.83	8.61
268	30	**15.4803**	14.54	9.23
269	15	19.6312	19.66	**11.09**
270	18	**14.0968**	11.13	7.99
271	26	**15.3338**	12.83	8.61
272	38	17.1834	**17.38**	10.26
273	23	**23.6493**	24.22	12.74
274	16	19.0316	**15.68**	9.64
275	31	**16.1934**	11.69	8.19
276	30	**20.3135**	20.23	11.29
277	30	**20.2301**	19.66	11.09
278	21	**22.1333**	19.66	11.09
279	23	**17.9749**	14.54	9.23

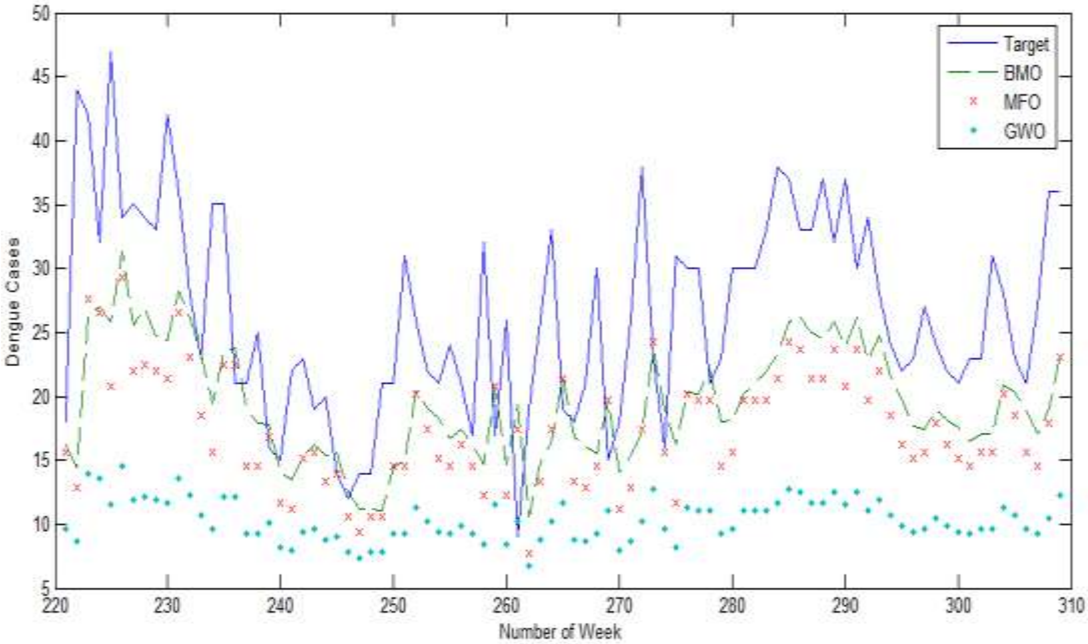


Figure 4- Target vs. BMO vs. MFO vs. GWO using normalized data

The convergence rate achieved by all the identified prediction algorithms are illustrated in Figure-5. Based on the figure, it is visualized that all the identified algorithms converged at the same value, however, the BMO is identified as the fastest.

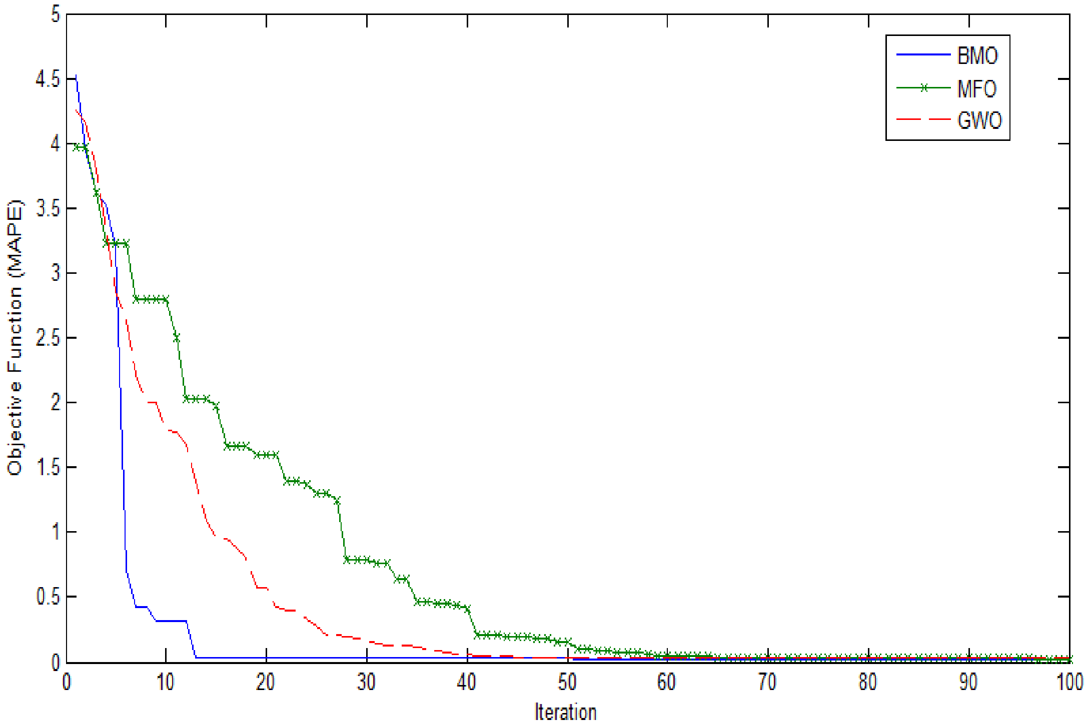


Figure 5- Convergence rate for BMO vs. MFO vs. GWO using normalized data

Meanwhile, results tabulated in Table-5 show the statistical level of the difference of the means between BMO prediction model and the other two identified prediction models viz. MFO and GWO are significant at 0.05% significance level.

Table 5- Significant Test for Dengue Outbreak Prediction using BMO, MFO and GWO

Data	Sig. (2 tailed)
BMO - MFO	.000
BMO - GWO	.000

Conclusion and future work

This study presents an application of BMO in dengue outbreak prediction. The case study covered the collected data from several districts in Peninsular Malaysia. The experiments were divided into two stages, where the first experiment was conducted using raw data while the latter using normalized data. The purpose is to find the best suite data for the proposed prediction model. The obtained results indicated that BMO performs better by using normalized data. Later, the BMO was compared against MFO and GWO which the results are in favour to BMO, relative to the identified metrics. By producing significant difference against MFO and GWO, the proposed BMO can benefit the interested parties in combatting dengue outbreak.

Acknowledgement

The authors thank the Faculty of Computing (http://www.ump.edu.my/ms/link/fakulti-komputeran) and Universiti Malaysia Pahang for their financial support (Grant No.RDU180366).

References

1. Hui Z. and Sanyang L. **2019**. Face Recognition based on Genetic Algorithm. *Journal of Visual Communication and Image Representation,* **58**: 495–502.
2. Haupt, R. L. and Haupt, S. E. **2004**. *Practical Genetic Algorithms*. Second ed. New Jersey: A John Wiley & Sons, Inc., Publication.
3. Jagat Kishore Pattanaik, Mousumi Basu, Deba Prasad D. **2018**. Improved real coded genetic algorithm for dynamic economic dispatch. *Journal of Electrical Systems and Information Technology*, **5**: 200–207.
4. Farnoosh Farhadian, Mohammad Mansour Riahi Kashani, Javad Rezazadeh, Reza Farahbakhsh, Kumbesan Sandrasegaran. **2019**. An Efficient IoT Cloud Energy Consumption based on Genetic Algorithm. *Digital Communication and Networks*.
5. Lee, C. K. H. **2018**. A review of applications of genetic algorithms in operation management. *Engineering Applications of Artificial Intelligence*,. **76**: 1–12.
6. Kennedy and Eberhart, R. Particle Swarm Optimization. **1995**. Proceedings of the IEEE International Conference on Neural Networks (ICNN): 1942–1948.
7. Dorigo, M. Swarm-Bots and Swarmanoid: Two Experiments in Embodied Swarm Intelligence. 2009. Proceedings of the IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies (WI-IAT): 2–3.
8. Mirjalili, S., Mirjalili, S. M. and Lewis, A. **2014**. Grey Wolf Optimizer. *Anvances in Engineering Software,* **69**: 46–61.
9. Mirjalili, S. **2015.** Moth-flame optimization Algorithm: A novel nature-inspired heuristic paradigm. *Knowledge-Based Systems*. **89**: 228–249
10. Seyedali Mirjalili, Amir H.Gandomi, Seyedeh Zahra Mirjalili, Shahrzad Saremi, Hossam Faris. **2017**. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. *Advances in Engineering Software*. **114**: 163–191.
11. Sulaiman, M. H, Mustaffa, Z, Saari, M. M, Daniyal, H, Musirin, I., and Daud, M. R. **2018** Barnacles Mating Optimizer: An Evolutionary Algorithm for Solving Optimization. Proceedings of the IEEE International Conference on Automatic Control & Intelligent Systems (I2CACIS 2018)
12. Shalini Y.K. and Sanjay K. **2017**. PSO-ANN based diagnostic model for the early detection of dengue disease. *New Horizons in Translational Medicine.* **4**(1–4): 1–8
13. Zuriani Mustaffa, Mohd Herwan Sulaiman, Mohamad Farhan Mohamad Mohsin, Yuhanis Yusof, Ferda Ernawan, Khairunnisa Amalina Mohd Rosli. **2019**. An Application of Hybrid Swarm Intelligence Algorithms for Dengue Outbreak Prediction. *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*. 731–735.
14. Zhiwei Xu, Hilary Bambrick, Laith Yakob, Gregor Devine, Francesca D.Frentiu, Rina Marina,

Pandji Wibawa Dhewantara, Roy Nusa, R. Tedjo Sasmono, Wenbiao Hu. **2019**. Using dengue epidemics and local weather in Bali, Indonesia to predict imported dengue in Australia," *Environmental Research.* **175**: 213–220

15. W. A. Er A. C. **2016**. Menangani wabak Denggi di Malaysia: Satu tinjauan kaedah rawatan dan pencegahan. *Malaysian Journal of Society and Space.* **12**(9): 56–68.

16. Vivek Jason Jayaraj, Richard Avoi, Navindran Gopalakrishnan, Dhesi Baha Raja, Yusri Umasa. **2019**. Developing a Dengue Prediction Model based on Climate in Tawau, Malaysia. *Acta Tropica.***197**: 1-7

17. Hani M. Aburas, Gultekin Cetiner, B., Murat Sari. **2010**. Dengue confirmed-cases prediction: A neural network model. *Expert Systems and Applications*. **37**: 4256–4260.

18. Zuriani Mustaffa, Mohd Herwan Sulaiman, Ferda Ernawan, Yuhanis Yusof and Mohamad Farhan Mohamad. **2018**. Dengue Outbreak Prediction: Hybrid Meta-Heuristic Model. Proceedings of the 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*:* 271–274.

19. Nor Azura Husin, Norwati Mustapha, Md. Nasir Sulaiman. **2012**. A hybrid model using genetic algorithm and neural network for predicting dengue outbreak. Proceedings of the 4th Conference on Data Mining and Optimization (DMO): 23–27.

20. Barazandeh, M., Davis, C. S., Neufeld, C.J., Coltman, D.W., and Palmer, A. R. 2**013**. Something Darwin didn't know about barnacles: spermcast mating in a common stalked species. Proceedings of the Royal Socienty B: Biological Sciences.

21. Yamaguchi, S., Charnov, E. L, Sawada, K and Yusa, Y. **2012**. Sexual Systems and Life History of Barnacles: A Theorethical Perspective. *Integrative and Comparative Biology.* **52**(3): 356–365

22. Mohd Herwan Sulaiman, Zuriani Mustaffa and Omar Aliman. **2019**. An Application of Barnacles Mating Optimizer for Solving Economic Dispatch Problems. IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT): 835–839.

23. Sukirty Jain, Sanyam Shukla, Rajesh Wadhvani. **2018**. Dynamic selection of normalization techniques using data complexity measures. *Expert Systems and Appicationsl.* **106**: 252-262