ISSN 2228-9860 eISSN 1906-9642 CODEN: ITJEA8

International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies

http://TuEngr.com

Optimized Deep Learning Procedure by Adaptive Parameters Based Genetic Algorithms for Determining Reservoir Inflow

Krotsuwan Phosuwan[1] and Panuwat Pinthong[1*]

[1] Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Thailand.
** (Tel: +066 858 317 983, s6102032910517@email.kmutnb.ac.th)*

Paper ID: 12A10P

Volume 12 Issue 10

Received 29 July 2021
Received in revised form 14 October 2021
Accepted 22 October 2020
Available online 06 August 2021

Keywords:
Rainfall-Runoff model; Reservoir Operation; Multilayer Perceptron; ANNs; AGA; APOGA; APGA; Kaeng Krachan reservoir; Phetchaburi river basin; Optimum deep learning parameters; Remaining Life Time (RLT); Water resource manangement; Reduced flood; Reservoir management.

Abstract

The determination of the reservoir inflow would be directly affected the efficiency of reservoir operation. Artificial intelligence techniques such as Artificial Neural Networks, Deep Learning (DL), and Genetic Algorithms (GAs) have been applied to many case studies of water resource management, for example, the determined relationship between rainfall and runoff, and rainfall forecast. DL has been successful for the rainfall-runoff model, but the performance of the model depends on its parameters that take more time-consuming for model development and is difficult to determine the optimum values. This paper presents the development of the Adaptive Parameters Based Genetic Algorithms (APGA) model to explore the optimum procedure of deep learning for reservoir inflow simulation for the Kaeng Krachan Reservoir and compare performance with the Adaptive Genetic Algorithm (AGA). The current study found that the mean absolute percentage error (MAPE) of the reservoir inflow from APGA was lower than AGA in all periods, so the optimum DL procedure from APGA outperforms AGA, while the DL layer architecture from APGA was more complex than AGA. In summary, APGA may be suitable for determining optimum DL procedure than AGA, but p_c and p_m parameters should be studied in the future.

Disciplinary: Civil Engineering & Technology (Hydrology), Computer Application.

Cite This Article:

Phosuwan, K., Pinthong, P. (2021). Optimized Deep Learning Procedure by Adaptive Parameters Based Genetic Algorithms for Determining Reservoir Inflow. *International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies, 12*(10), 12A10P, 1-11. http://TUENGR.COM/V12/12A10P.pdf DOI: 10.14456/ITJEMAST.2021.205

1 Introduction

The accurate determination of the reservoir inflow directly affects the efficiency of reservoir operation, such as increasing the capacity of flood protection and reducing water shortage in the dry season [1].

In recent years, many researchers have developed and utilized methods to study the rainfall-runoff relationship. One is Artificial intelligence techniques, such as Artificial Neural Networks (ANNs), Machine Learning (ML), Deep Learning (DL), and Genetic Algorithms (GAs).

DL has been applied to many case studies of water resource management problems. In [2], ANNs were used to study the relationship between rainfall and runoff in a rural drainage area in southern England that has a watershed area of about 30 km^2. The result found that the impact of the complex of the networks may affect the model performance, and less complex network generalizes better performance than more complex networks. In addition, ANNs could simulate the runoff accurately in this case study. Unlike [3], the results obtained that the performance of prediction from multilayer architecture model had better performance than the single-layer architecture model.

DL has many parameters such as weights, bias, number of hidden layers, activation functions, and hyperparameters that affect the accuracy of the model [4]; consequently, some researchers have been applied optimization techniques, for example, Genetic Algorithms (GAs) [5], and Particle Swarm Optimization (PSO) [6] for tuning DL parameters. GAs were explained in [7] are a stochastic search technique that has been applied and successful in many real-world optimization problems to determine optimum variables in the solution, and their three important parameters are population size, probability of crossover, and probability of mutation. One study by [8] developed the GA-DLNN hybrid model adapted from GAs and DL to predict the bearing capacity of the driven pile, the result obtained that the GA-DLNN hybrid model could search for the optimum DL parameters for the prediction process.

Following the first of developing GAs, many researchers have emphasized the improving performance of GAs. For example, [9] proposed the Adaptive Population Pool Size Based Genetic Algorithm (APOGA). In this study, the population pool size of GAs was modified to play a role in objective fitness. Moreover, the result showed that APOGA surpasses standard GAs. In [10] applied the Adaptive Genetic Algorithm (AGA) of which parameters consist of the probability of mutation and probability of crossover that were modified automatically in each generation depend on population fitness. As the result, the AGA used less execution time and better performance than standard GAs.

This study develops the Adaptive Parameters Based Genetic Algorithms (APGA) model that self-tuning population size, probability of mutation, and probability of crossover parameters, and scrutinize its performance by applying it to explore optimum parameters of deep learning for reservoir inflow simulation; furthermore, the statistical model will be used to evaluate the model performance and compare between APGA and AGA model.

2 Materials and Methodology

2.1 Study Area

The selected case study in this research at the Kaeng Krachan Reservoir locates in the Phetchaburi River Basin that lies between 12°30’ and 13°30’ north of the equator and at eastern longitude from 99° to 100°15’ east in Southwest of Thailand (Figure 1). With a watershed area of 2,210 km^2, the mean annual rainfall is 1,046 mm per year and the average reservoir inflow is about 917 million cubic meters (MCM) per year. This reservoir is the most important tool for supply water to the demand area and reduces floods at the downstream area of the Phetchaburi River Basin. Table 1 shows some characteristics of the reservoir, it can be seen that the mean annual inflow was higher than the reservoir capacity.

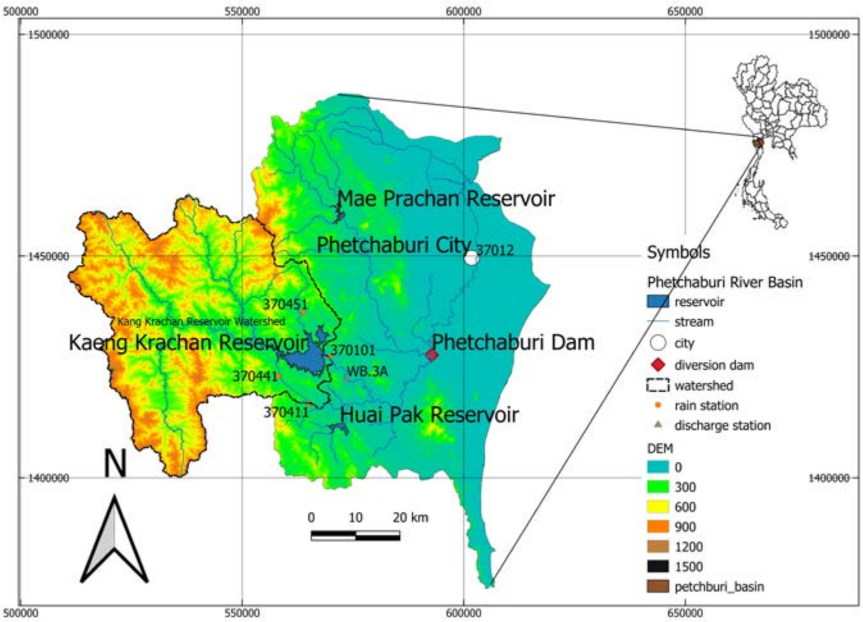

Figure 1: Location of the Kaeng Krachan Reservoir and its watershed in the Phetchaburi River Basin of Thailand.

Table 1: Characteristics of the Kaeng Krachan Reservoir

Parameters	Kaeng Krachan Reservoir
Watershed area (km^2)	2,210.00
Maximum storage volume (MCM)	895.30
Normal storage volume (MCM)	710.00
Dead storage volume (MCM)	65.00
Mean annual inflow (MCM)	917.00
Mean annual rainfall (mm)	1,046.00

2.2 Data Collection

The rain gauge datasets and reservoir inflow datasets are the historical daily time series obtained from the Royal Irrigation Department. Four selected ground base rainfall stations in this study were 370451, 370101, 370441, and 370411. In addition, reservoir inflow datasets were calculated by the water balance of the Kaeng Krachan Reservoir. Figure 1 above provides the location of the selected rainfall station and the Kaeng Krachan Reservoir. Table 2 shows some statistical information about rainfall and reservoir inflow datasets. It can be seen from its data that the highest rainfall was in the training period, and the highest inflow was in the testing periods.

Table 2: List of the datasets statistical information.

Datasets		Statistical Parameters			Period
		μ	maximum	minimum	
rainfall (mm)	training	2.35	154.03	0.00	2006-2017
	testing	2.63	75.60	0.00	2018-2019
	application	3.37	97.55	0.00	2020
inflow (MCM)	training	2.47	70.33	0.00	2006-2017
	testing	3.91	110.69	0.00	2018-2019
	application	2.06	25.96	0.00	2020

Because the previous data of rainfall and reservoir inflow variables may be used as input for the rainfall-runoff (or reservoir inflow) model [2], the first five previous rainfall and reservoir inflow have the highest correlation with reservoir inflow were investigated. The results of the correlation analysis were summarised in Table 3 so that antecedent daily rainfall from 1 to 5 days $\{P_{t-1} \dots P_{t-5}\}$ and antecedent daily inflow from 1 to 5 days $\{I_{t-1} \dots I_{t-5}\}$ were combined to create 25 data sets, then were input to DL model and selected for the best performance of simulation inflow (I_t) model. The inflow at time step t (I_t) was made from antecedent daily rainfall and reservoir inflow is given by

$$I_t = f\{I_{t-1}, I_{t-2} \dots I_{t-5}, P_{t-1}, P_{t-2} \dots P_{t-5}\} \quad (1).$$

Table 3: The first five previous rainfall and reservoir inflow have the highest correlation with reservoir inflow; sorted by correlation descending

No	Antecedent daily rainfall	Antecedent daily inflow
1	P_{t-2}	I_{t-1}
2	P_{t-3}	I_{t-2}
3	P_{t-1}	I_{t-3}
4	P_{t-4}	I_{t-4}
5	P_{t-5}	I_{t-5}

The best input dataset was selected by the lowest mean absolute error (MAE) of the DL model. After that, it was used as input for the DL model to compared the performance between APGA and AGA models.

2.3 Deep Learning

TensorFlow library [11] was used for developing the DL model, and Multilayer Perceptron (MLP) as part of the deep learning method was chosen for simulating reservoir inflow in this research.

2.4 Adaptive Parameters Based Genetic Algorithms: APGA

In this study, the Adaptive Parameters Based Genetic Algorithms: APGA was developed and adapted from Adaptive Population Pool Size Based Genetic Algorithm: APOGA in [9] and Adaptive Genetic Algorithm: AGA in [10]; as a result, APGA adjusts Genetic Algorithms parameters consists of population size, probability of crossover, and probability of mutation automatically. The objective function of the APGA is to minimize the mean absolute error (MAE) that is a measure of errors between simulate reservoir inflow from DL model versus observation datasets, and the five parameters of the DL model procedure were selected as variables include layer architecture, layer activation functions, learning rate, epochs, and batch size. The flow chart of APGA is presented in Figure 2, with the preliminary details as below.

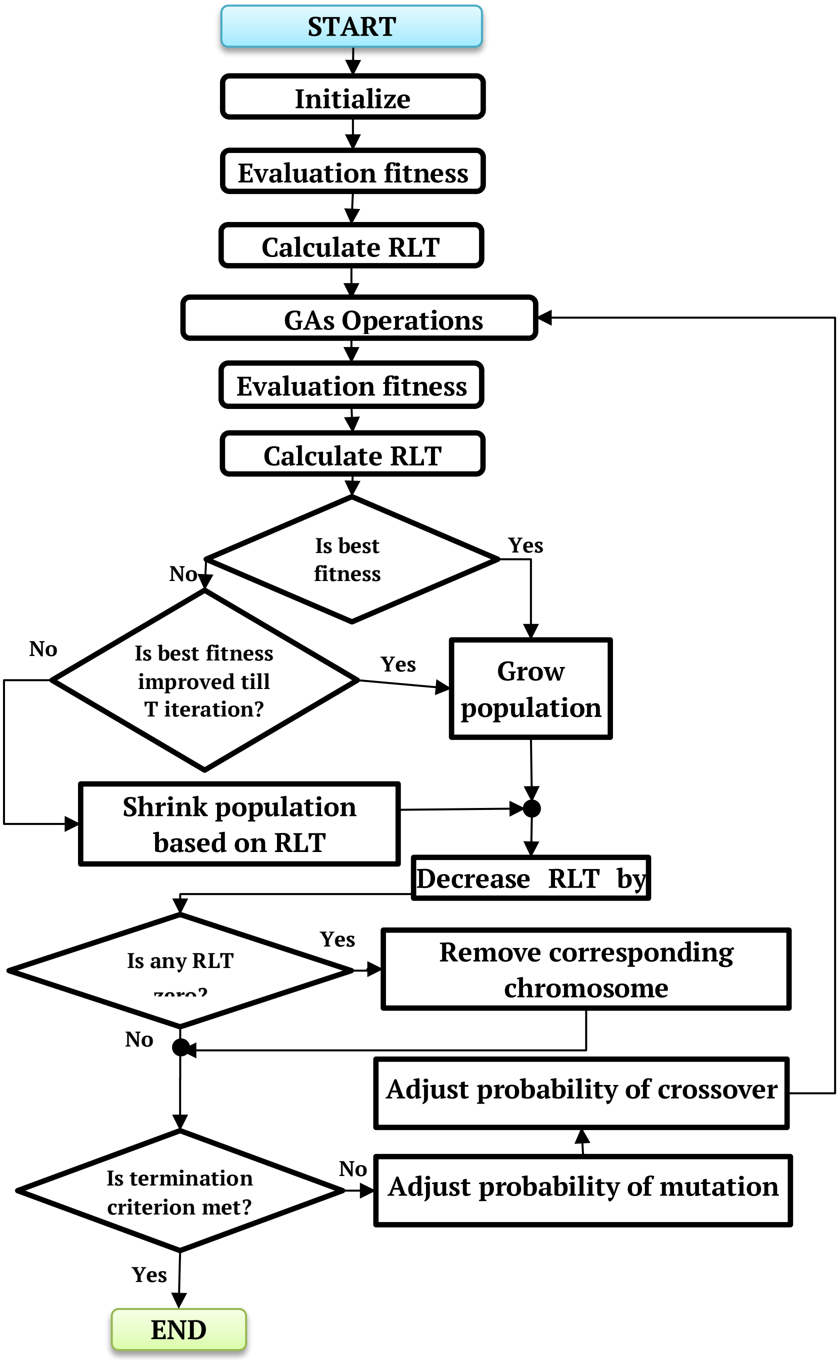

Figure 2: Flowchart of APGA. Adapted from [9].

The Remaining Life Time (RLT) is a lifetime measure for each chromosome calculated by the mathematical formulation in Equation 2 [9].

$$RLT(i) = \begin{cases} LT^{min} + \eta \frac{F^{worst} - F(i)}{F^{worst} - F^{avg}}; if\ F(i) > F^{avg} \\ \frac{1}{2}\left(LT^{min} + LT^{max}\right) + \eta \frac{F^{avg} - fitness(i)}{F^{avg} - F^{best}}; if\ F(i) < F^{avg} \end{cases} \quad (2)$$

Where $\eta = \frac{1}{2}\left(LT^{max} - LT^{min}\right)$; LT^{max} and LT^{min} are the maximum and minimum lifetimes of the chromosome in the population pool, respectively. In each generation, RLT is monitored and calculated new value. Once the RLT of the chromosome attains zeros, the chromosome is removed from the population pool.

As in [9], the RLT of the new chromosome is calculated using Equation (2). Depending on the RLT of the entire population pool and the fitness values, the resizing arises. For resizing, the population size is grown at a rate as

$$G = \alpha \; x \; (I^{max} - I) x \frac{F^{newbest} - F^{oldbest}}{F^{initialbest}} \quad (3).$$

In Equation (3), α is a random real value that is greater than or equal to 0.0 and less than 1.0, G, $F_{newbest}$, $F_{initialbest}$, I_{max}, and I are growth size, best fitness value of current iteration, best fitness value of previous iteration, initial best fitness, the maximum number of generation and current iteration number, respectively.

$$P_c = \begin{cases} k_1 (f_{min} - f')/(\bar{f} - f_{min}) & ; f' \le f_{min} \\ k_3 & ; f' > f_{min} \end{cases} \quad (4)$$

$$P_m = \begin{cases} k_2 (f - f_{min})/(\bar{f} - f_{min}) & ; f \ge f_{min} \\ k_4 & ; f < f_{min} \end{cases} \quad (5)$$

The probability of crossover (p_c) and probability of mutation (p_m) are calculated by Equations (4) and (5), which consist of average fitness value ($\bar{f}$) of the population, minimum fitness value f_{min} of the population and the best fitness (f') presents in Equations (4) and (5). As analyzed by [10], the values of k_4 and k_2 were assigned to 0.5, while k_1 and k_3 values were assigned to 1.0.

2.5 Evaluation Criteria for Model Comparison

The mean absolute percentage error (MAPE) which was described in [12] was chosen as the statistical tool for measuring and comparing model performance.

3 Results and Discussion

3.1 Best Input Set for DL Model

Because the input data of DL affected the model performance, the antecedent rainfall and reservoir inflow datasets were combined to created twenty-five datasets and used as input to APGA for optimum DL procedure. After that, the best result of all models that have the lowest MAE was selected, as a result, the input group I_{t-1}, I_{t-2}, I_{t-3}, I_{t-4}, P_{t-1}, P_{t-2}, P_{t-3}, P_{t-4}, and P_{t-5} affected the lowest MAE of APGA and was selected as input for APGA and AGA model for comparing model performance.

3.2 Sensitivity Analysis for Population Size of AGA Model

The sensitivity analysis for population size is an important process as the consequence of the population size parameter is directly affected AGA model performance. AGA has a process for automatically adjusting the probability of crossover and probability of mutation in each generation; nevertheless, population size is fixed for all generations. In this process, the different values of population sizes (population size = 20 to 100 in steps of 10) were used for sensitivity analysis for population size were performed. In sum, while the population size equaled to 90, AGA obtained an optimum solution that MAE was at 0.636 MCM, as the result shown in Figure 3.

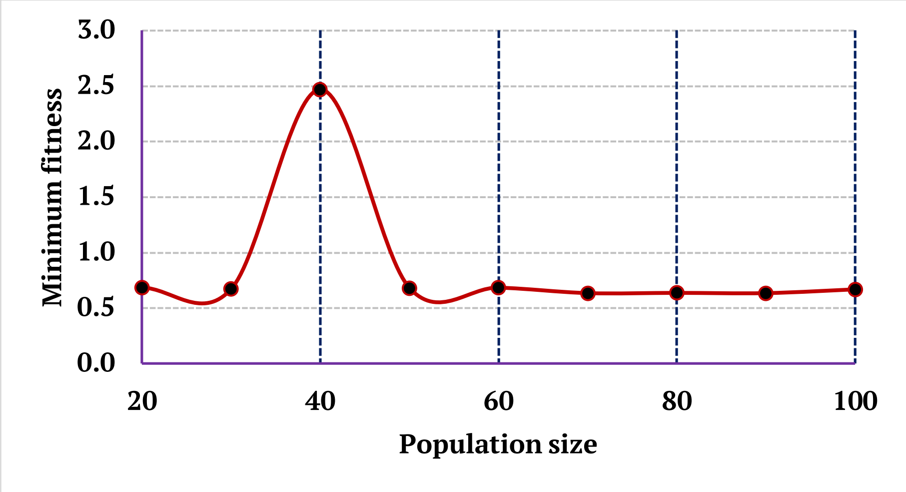

Figure 3: Sensitivity analysis for the population size of AGA model

3.3 Model Comparison and Reservoir Inflow Simulation

The results of the DL procedure of the APGA and AGA model were shown in Table 4, and they were discovered that the layer architecture of the DL model of the APGA model was more complex than the AGA model.

Table 4: Optimum Deep Learning parameters

Parameters	APGA	AGA
Layer architecture	9, 440, 358, 52, 183, 283, 435, 280, 396, 49, 243, 73, 109, 405, 1	9, 136, 201, 213, 188, 215, 101, 1
Layer activators	relu, relu, relu, tanh, tanh, relu, relu, tanh, elu, tanh, relu, -, tanh, elu, tanh	tanh, elu, relu, elu, relu, relu, tanh
Learning rate	0.0001	0.0007993
Epochs	17	15
Batch size	116	559

Table 5: Performances of APGA and AGA models

Statistic	Model	Training	Testing	Application
MAPE (%)	APGA	16.73	15.40	16.88
	AGA	17.66	17.08	18.04
R^2	APGA	0.945	0.973	0.922
	AGA	0.883	0.953	0.825

While the study was in process of comparing model performance, the MAPE was used and presented in Table 5. It can be seen that the MAPE of the APGA model was lower than the AGA model, and R^2 of the APGA model was higher than the AGA model at the training, testing, and application periods. The most striking result to emerge from Table 5 is that the APGA model may outperform the AGA model. The comparison of the fitness values of the APGA and AGA model in each generation, as shown in Figure 4, indicated that the fitness of both models dramatically decreases after the initial population to the third generation, and the APGA model had lower fitness than the AGA model in all generations. From this figure, it can be seen that the APGA model may have more performance than the AGA model.

The percent errors in training, testing, and application periods were shown in Figures 5, 6, and 7, respectively. In the training period, the APGA model had percent errors of inflow between 0.00 to 410.50 percent, and AGA had percent errors between 0.01 to 336.37 percent. In the testing period, the APGA model had percent errors between 0.04 to 236.26 percent, and AGA had percent errors between 0.19 to 354.64 percent. In addition, in the application period, the APGA model had percent errors between 0.31-94.47%, and AGA had percent errors between 0.07-162.30%.

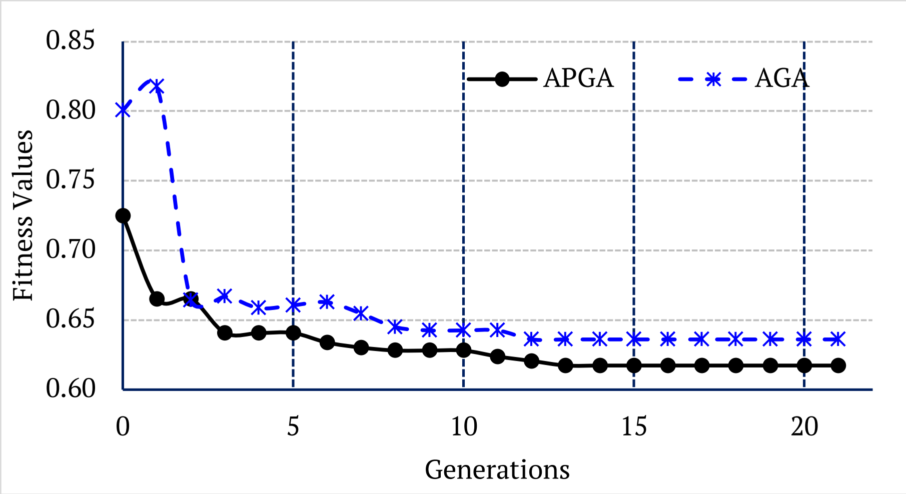

Figure 4: Comparison of the fitness values of the APGA and AGA models.

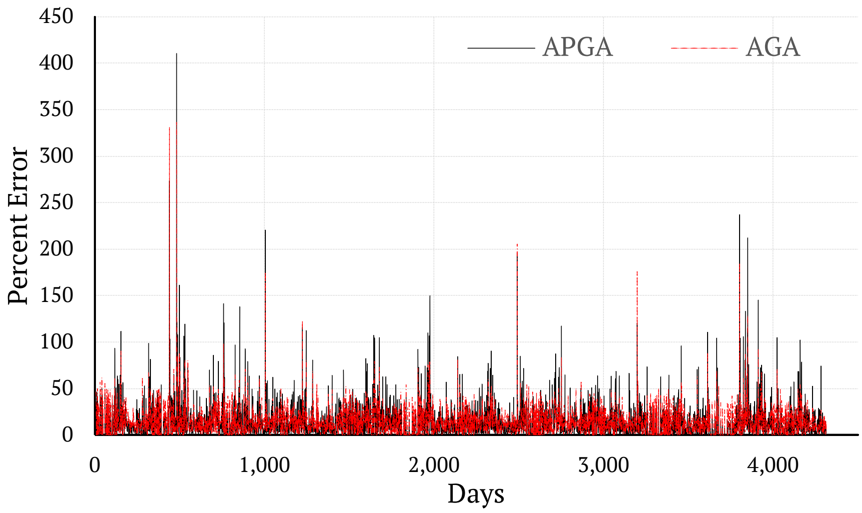

Figure 5: The comparison of percent error from APGA and AGA models during the training period

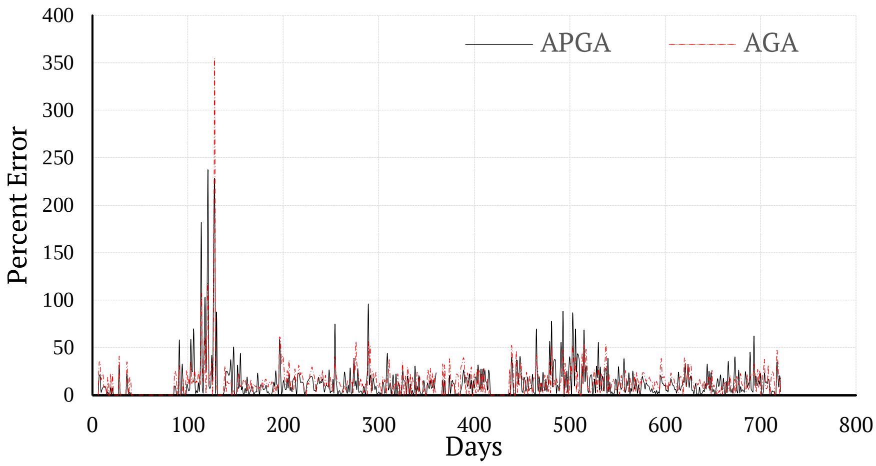

Figure 6: The comparison of percent error from APGA and AGA models during the testing period.

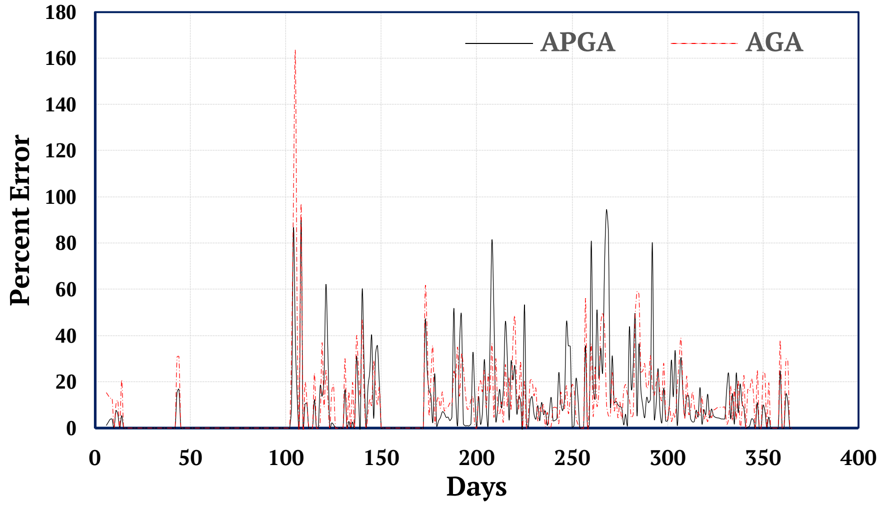

Figure 7: The comparison of percent error from APGA and AGA models during the application period

Figures 8, 9, and 10 are presented the comparison between observed reservoir inflow, simulation results of the APGA and AGP models during the training, testing, and application period, respectively. In the training period, the maximum inflow of observed data, APGA, and AGA models were 70.33, 44.86, and 55.94 MCM, respectively. In the testing period, the maximum inflow of observed data, APGA, and AGA models were 110.69, 46.84, and 43.17 MCM, respectively. In the application period, the maximum inflow of observed data, APGA, and AGA models were 25.95, 16.52, and 19.86 MCM, respectively. In general, both models satisfactorily simulated the reservoir inflow, but they were inclined to underestimate the peak-flow events.

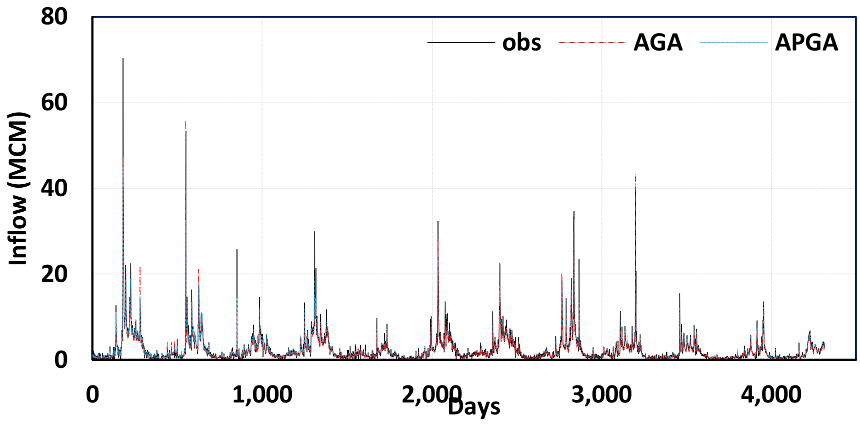

Figure 8: The comparison of inflow from observation data, APGA, and AGA models during the training period.

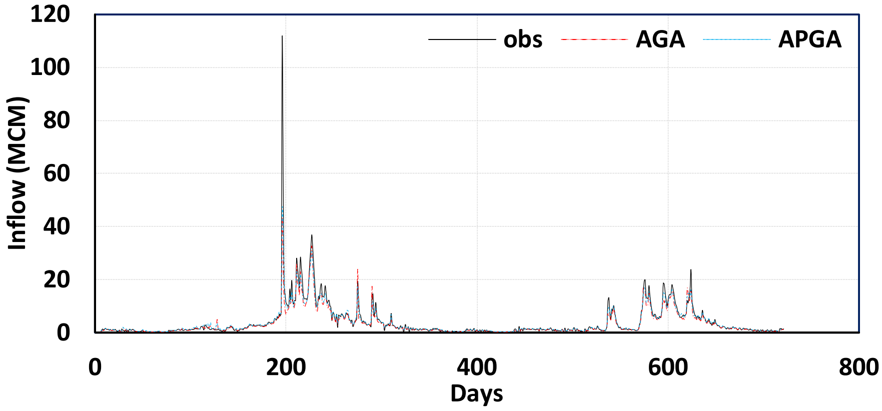

Figure 9: The comparison of inflow from observation data, APGA, and AGA models during the testing period.

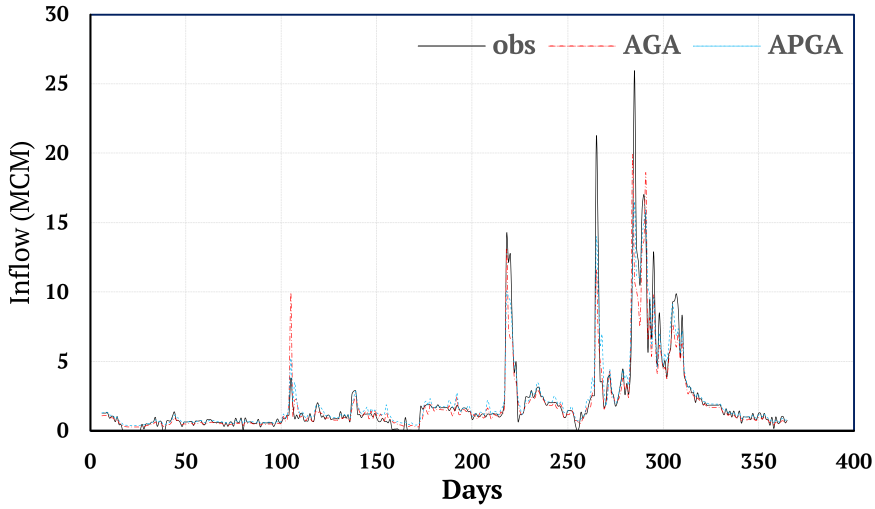

Figure 10: The comparison of inflow from observation data, APGA, and AGA models during the application period

4 Conclusion

In this study, Adaptive Parameters Based Genetic Algorithms (APGA) was developed by adapting from Adaptive Population Pool Size Based Genetic Algorithm (APOGA) and Adaptive Probabilities of Crossover and Mutation Genetic Algorithm (AGA) for tuning DL parameters for determining reservoir inflow; consequently, its parameters, compose of the probability of crossover, probability of mutation, and population size, were changed automatically in each generation. APGA model has a lower fitness value and MAPE of the DL model than the AGA model. The results of this study indicated that the APGA model outperforms the AGA model for tuning deep learning parameters. At this point, the accurate inflow simulation from DL that was tuned by the APGA model may improve the efficiency of reservoir operation and reduce the impact of flooding. In addition, it may be suitable for applying to other real-world optimization problems. In future research, APGA will be applied to optimize multiple reservoir releases in the Phetchaburi River Basin to reduce flood and drought in the downstream area.

5 Availability of Data and Material

Data can be made available by contacting the corresponding author.

6 References

[1] S. Yang and T. Yang, "Uncertainty Assessment: Reservoir Inflow Forecasting with Ensemble Precipitation Forecasts and HEC-HMS," Advance in Meteorology, pp. 1-11, 2014. https://doi.org/10.1155/2014/581756

[2] A. W. Minns and M. J. Hall, "Artificial Neural Networks as Rainfall-Runoff Models," Hydrological Sciences journal, pp. 399-417, 1996. DOI: 10.1080/02626669609491511

[3] J. Heaton, "Deep Learning and Neural Networks," Heaton Research Publisher Vol.3, 2015.

[4] S. Afaq and S. Rao, "Significance Of Epochs On Training A Neural Network," International Journal Of Scientific & Technology, VOLUME 9, pp. 485-488, 2020.

[5] D., Rivero, Dorado J., Fernández-Blanco E., Pazos A., "A Genetic Algorithm for ANN Design,

Training and Simplification,". In: Cabestany J., Sandoval F., Prieto A., Corchado J.M. (eds) Bio-Inspired Systems: Computational and Ambient Intelligence. IWANN 2009. *Lecture Notes in Computer Science*, vol 5517. Springer, Berlin, Heidelberg, 2009. https://doi.org/10.1007/978-3-642-02478-8_49

[6] G. A. Beatriz and V. A. Roberto, "Designing Artificial Neural Networks Using Particle Swarm Optimization Algorithms," Computational Intelligence and Neuroscience, 2015. https://doi.org/10.1155/2015/369298

[7] S. N. Sivanandam and S. N. Deepa, " Introduction to Genetic Algorithms", New York: Springer, 2008.

[8] T. A. Pham, V. Q. Tran and H.-L. T. Vu, "Design Deep Neural Network Architecture Using a Genetic Algorithm for Estimation of Pile Bearing Capacity," LoS ONE, pp. 1-25, 2020. https://doi.org/10.1371/journal.pone.0243030

[9] B.R.Rajakumara and A. George, "APOGA: An Adaptive Population Pool Size Based Genetic Algorithm," 2013.

[10] A. Laoufi, S. Hadjeri and A. Hazzab, "Adaptive Probabilities of Crossover and Mutation in Genetic Algorithms for power economic dispatch," International Journal of Applied Engineering Research, vol. 1(3), pp. 393-408, 2006.

[11] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner and P. Tucker, "TensorFlow: A System for Large-Scale Machine Learning," in 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI ’16), Savannah, GA, USA, 2016.

[12] P. M. Swamidass "MAPE (mean absolute percentage error) Mean Absolute Percentage Error (MAPE)." Encyclopedia of Production and Manufacturing Management, 462-462, 2000. DOI: 10.1007/1-4020-0612-8_580

Krotsuwan Phosuwan is a student at the Department of Teacher Training in Civil Engineering, King Mongkut’s University of Technology North Bangkok, Thailand. He got a Bachelor’s degree in Irrigation Engineering from Kasetsart University, Thailand. His research includes Water Resource Management and Computer Application Development.

Dr.Panuwat Pinthong is a Faculty member at the Department of Teacher Training in Civil Engineering, and Head of Center for Water Engineering and Infrastructures Research King Mongkut’s University of Technology North Bangkok. He got a Ph.D. (Water Engineering and Management) from the Asian Institute of Technology, Thailand.