© Universiti Tun Hussein Onn Malaysia Publisher's Office

IJIE

http://penerbit.uthm.edu.my/ojs/index.php/ijie
ISSN : 2229-838X e-ISSN : 2600-7916

The International Journal of Integrated Engineering

Metaheuristic Algorithms to Enhance the Performance of a Feedforward Neural Network in Addressing Missing Hourly Precipitation

Wai Yan Lai[1], King Kuok Kuok[1*], Shirley Gato-Trinidad[2], Md. Rezaur Rahman[3], Muhammad Khusairy Bakri[3]

[1]Faculty of Engineering, Computing and Science,
Swinburne University of Technology, 93350 Kuching, Sarawak, MALAYSIA

[2]Faculty of Engineering, Computing and Science,
Swinburne University of Technology, Hawthorn Victoria 3122, AUSTRALIA

[3]Faculty of Engineering,
Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, MALAYSIA

*Corresponding Author

DOI: https://doi.org/10.30880/ijie.2023.15.01.025
Received 23 April 2022; Accepted 20 December 2022; Available online 31 March 2023

Abstract: This research study investigates the implementation of three metaheuristic algorithms, namely, Grey Wolf Optimizer (GWO), Multi-Verse Optimizer (MVO), and Moth-Flame Optimisation (MFO), for coupling with a Feedforward Neural Network (FNN) in addressing missing hourly rainfall observations, while overcoming the limitation of conventional training algorithm of artificial neural network that often traps in local optima. The proposed GWOFNN, MVOFNN, and MFOFNN were compared against the conventional Levenberg Marquardt Feedforward Neural Network (LMFNN) in addressing the artificially introduced missing hourly rainfall records of Kuching Third Mile Station. The findings show that the proposed approaches are superior to LMFNN in predicting the 20% hourly rainfall observations in terms of mean absolute error (*MAE*) and coefficient of correlation (*r*). The best performance ANN model is GWOFNN, followed with MVOFNN, MFOFNN and lastly LMFNN.

Keywords: Grey Wolf Optimizer (GWO), Multi-Verse Optimizer (MVO), Moth-Flame Optimization (MFO), artificial neural network, missing hourly rainfall observations

1. Introduction

Rainfall is one of the most important variables that often utilized and included for further hydrological and climatological modelling, analysis and simulation studies [1]. It is required to have long term, consistent, and high resolution of climate variables that include the sub-daily rainfall observations for executing them accurately [2]. However, the occurrences of missing rainfall observations are unavoidable, mostly due to human errors in data management, natural disasters, and machinery defects on site when measuring and collecting the rainfall data. There are three conventional approaches for addressing missing data observations in various fields of research as the conventional data pre-processing approaches, and they are hot-deck imputation, listwise deletion, and zero imputation [3], [4]. Hot-deck imputation is currently adopted in Malaysia to address the missing rainfall observations by substituting the missing entries with the entries that were measured simultaneously from neighboring rainfall measurements [5].

**Corresponding author: kkuok@swinburne.edu.my*

However, the conventional approaches are not reliable, and they are not scientifically supported to be accurate and feasible in all the scenarios. Thus, data imputation approaches are encouraged by researchers to deal with the missing variables.

Artificial Neural Network (ANN) is one of the most popular Artificial Intelligence (AI) based empirical methods in predicting hydrological and climate variables [6]. The reason that ANNs are popular for hydrological predictions is due to the ability to account for the non-linear relationships of hydrological and climatological variables, and they do not rely on high data demand in order to achieve reliable and accurate predictions, which make them much more preferrable than other physical based and conceptual models [7], [8]. Back-Propagation Neural Network (BPNN) had successfully applied for estimating the monthly rainfall runoff of the Hub River in Pakistan, and reliable estimations could be obtained via BPNN [9]. Wavelet Neural Network model (WNN) was found outperformed than the conventional Multilayer Perceptron (MLP) for predicting the monthly rainfall of the Darjeeling rain gauge station [10]. On the other hand, ANNs are also proven to be feasible in addressing missing rainfall data [11]. Through the performed rainfall imputation study, the Levenberg-Marquardt (LM) algorithm was found to be superior in training the Feedforward Neural Network (FNN) when compared against the Conjugate Gradient Fletcher–Reeves update (CGF) algorithm and Broyden–Fletcher–Goldforb–Shanno (BFG) algorithm [11]. In addition to that matter, the LM algorithm is also highly recommended by MATLAB software as the supervised algorithm for training ANN due to its superior performance and time efficiency [12]. Apart from that, other AI-based empirical imputation models such as K-Nearest Neighbor (KNN) [1], [13], Sequential K-Nearest Neighbor (SKNN) [14], Regularized Expectation Maximization (RegEM) [15], Bayesian Principal Component Analysis (BPCA) [16], Expectation Maximization (EM) [17], and probabilistic principal component analysis (PPCA) [18] are also proven to be reliable in addressing missing hydrological and climatological variables.

Although ANNs are proven to be robust in various hydrological and climatological predictions, the conventional training algorithms of ANNs often traps in local optima, which it fails to achieve the global optimum hence limiting the prediction accuracy [19], [20]. In other words, the solution set in terms of Weights (W) and Biases (B) trained by the conventional training algorithms during the training stage is not feasible to be applied in the testing stage to make the predictions. In this case, metaheuristic algorithms can be applied to replace the conventional training algorithms of ANNs for local optima avoidance, which are simple and flexible for various optimization attempts [21]. Metaheuristic algorithms are designed by referring to the inspiration of natural systems, where they are mathematically formulated and developed based on the concept of evolution in nature, physics laws, or behavior of living organisms [22]. It is currently a popular approach to be coupled with ANNs to enhance the performance of the conventional ANNs. Grey Wolf Optimizer (GWO) algorithm was applied in training MLP for classifying Exclusive or (XOR), balloon, iris, breast cancer, and heart datasets [23]. It was found that the GWO algorithm outperformed the Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Ant Colony Optimization (ACO), Evolution Strategy (ES), and Population-based Incremental Learning (PBIL) in training MLP for the classification tasks. Similarly, Moth-flame Optimization (MFO) algorithm developed by Mirjalili was used to train MLP networks for solving classification problem [24], [25]. A competitive result can be observed by the authors when comparing the MFO algorithm against PSO, GA, ACO, and ES optimization approach in training MLP for classifying five different datasets obtained from the University of California at Irvine (UCI) Machine Learning Repository. On the other hand, hybrid approach of ANN with the Multi-Verse Optimizer (MVO) algorithm was adopted to predict the streamflow [26]. The findings showed that the MVO-ANN model is superior to the other ANN models trained by PSO and BP algorithms. The dragonfly algorithm (DA) medical prediction and classification tasks. It was found that the DA-ANN approach could outperform the existing classification approaches such as Bayes Network training with hill climbing algorithm, Naive Bayes, and KNN classifier.

From the review above, it shows that replacing the conventional training algorithms of ANNs with metaheuristic algorithms seems to be a plausible and reliable solution to enhance the prediction performance of ANNs. However, there is a lack of hourly imputation studies, especially the hourly rainfall imputation studies, where the recent rainfall imputation studies mostly focused on addressing annually, monthly, and daily scaled rainfall observations [27]. Thus, it inspired this study to adopt the GWO, MVO, and MFO algorithm to be coupled with FNN for addressing missing hourly rainfall observations. To the best of the authors' knowledge that the selected metaheuristic algorithms are yet applied together with FNN for addressing any missing sub-daily hydrological variables, especially for missing hourly rainfall observations. Hence, a case study was formulated and executed in this study for exploring the feasibility and robustness of the GWOFNN, MVOFNN, and MFOFNN in addressing missing hourly rainfall observations. By accounting and considering the aforementioned issues, the objectives of this study are outlined as listed below:

a) To predict the hourly missing rainfall observations with the formulated hybrid approaches of the selected metaheuristic algorithms with FNN (GWOFNN, MVOFNN, and MFOFNN),
b) To compare the performance of GWOFNN, MVOFNN, and MFOFNN against an existing conventional ANN, LMFNN,
c) To evaluate the reliability and feasibility of GWOFNN, MVOFNN, and MFOFNN in addressing missing hourly rainfall observations.